AFRL-OSR-VA-TR-2013-0343

DEVELOPMENT AND CHARACTERIZATION OF NOVEL BIOLUMINECENT SYSTEMS

Bruce Branchini
Connecticut College

July 2013
Final Report

DISTRIBUTION A: Approved for public release.

AIR FORCE RESEARCH LABORATORY
AF OFFICE OF SCIENTIFIC RESEARCH (AFOSR)
ARLINGTON, VIRGINIA 22203
AIR FORCE MATERIEL COMMAND

REPORT DOCUMENTATION PAGE

OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. **PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.**

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
05-03-2013	Final Performance Report	01-05-2010 — 30-04-2013

4. TITLE AND SUBTITLE
Development and Characterization of Novel Bioluminescent Systems

5a. CONTRACT NUMBER

5b. GRANT NUMBER
FA9550-10-1-0174

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Bruce R. Branchini

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Connecticut College
270 Mohegan Avenue
New London, CT 06320-4150
860 439-2081

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
AFOSR
875 N. Randolf Street
Arlington, VA 22203

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
AFRL-OSR-VA-TR-2013-0343

12. DISTRIBUTION / AVAILABILITY STATEMENT
A-Approved for public release

13. SUPPLEMENTARY NOTES

14. ABSTRACT

The overall objective of this project was to discover, design and demonstrate the feasibility of bioluminescent materials for use in marking, tagging, and anti-tamper applications.(**1**) The technology for near-Infrared (nIR) dye labeling and (BRET) activation of a firefly luciferase were transferred to Rajesh Naik's lab at AFRL.(**2**) The construction and evaluation of novel protease substrates (luciferase fusion proteins emitting nIR light) for caspase and the two blood clotting factors was completed.(**3**) Novel nIR emitting quantum dots/rods were made using luciferases in collaboration with Mathew Maye, Syracuse University. (**4**) A firefly luciferin benzothiophene-containing analog with glow kinetics and 523 nm emission was discovered. (**5**) Analyses of luminescent materials from *C. variopedatus* provided by the Deheyn lab revealed the presence of riboflavin.(**6**) A firefly luciferase variant was intramolecularly cross-linked trapping it in a C-domain rotated previously undocumented conformation. (**7**) With A. Gulick (Hauptman Woodward Institute) luciferase crystal structures were determined providing insight into the mechanism of the oxidative light reaction.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			19b. TELEPHONE NUMBER *(include area code)* 860 439-2479

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39.18

FINAL REPORT

BRUCE R. BRANCHINI
(05-01-2010 to 04-30-2013)
AFOSR FA9550-10-1-0174
May 03, 2013

I. OBJECTIVES AND SIGNIFICANCE

The **overall objective** is to discover, design and demonstrate the feasibility of bioluminescent materials for use in marking, tagging, and anti-tamper applications. Bioluminescent materials are non-toxic and biodegradable and can be incorporated into landing zone markers, anti-tamper systems, perimeter security systems, friend versus foe marking systems, etc. The bioluminescent materials only emit after proper activation, not spontaneously. Furthermore, their emission characteristics may be modified so that their signals are bright and decay after a few seconds or they can be made to last for 30 minutes to an hour. Moreover, the color of the light signal can be varied over the visible range and into the near Infrared (nIR) region by the bioluminescence energy transfer (BRET) process. Bioluminescent coatings may also be used in anti-tamper and other security-related applications.

The three major objectives investigated during the grant period were:

1. To develop firefly luciferases as efficient sources of nIR radiation.

2. *To develop and characterize new bioluminescent systems.*

3. *To further investigate the domain alternation mechanism in P. pyralis firefly luciferase (Luc) bioluminescence using site-directed mutagenesis, time-resolved fluorescence spectroscopy and crystallography.*

II. Summary of Significant Work Accomplished and Publications

There were 7 major accomplishments completed with this grant. They are presented below organized according to the three major objectives.

OBJECTIVE 1. To develop firefly luciferases as efficient sources of nIR radiation.

Publications:

1. "Red-Emitting Luciferases for Bioluminescence Reporter and Imaging Applications," B. R. Branchini, D. M. Ablamsky, A. L. Davis, T. L. Southworth, B. Butler, F. Fan, A. P. Jathoul and M. A. Pule, *Analytical Biochemistry* 396 290-297 (2010).
2. "Chemically Modified Firefly Luciferase Is an Efficient Source of Near-Infrared Light," B. R. Branchini, D. M. Ablamsky, and J. C. Rosenberg, *Bioconjugate Chemistry* 21 (11), 2023-2030 (2010).
3. "Biotechnological Improvements of Bioluminescent Systems," K. T. Hamorsky, E. Dikici, C. M. Ensor, S. Daunert, A. L. Davis and B. R. Branchini. In *Chemiluminescence and Bioluminescence Past, Present and Future*, edited by A. Roda, RSC Publishing, pp. 443-487, 2010 (ISBN 978-1-84755-812-1).
4. "A Portable Bioluminescence Engineered Cell-Based Biosensor for On-Site Applications," A. Roda, L. Cevenini E. Michelini and B. R. Branchini, *Biosensors and Bioelectronics* 26: 3647-3653 (2011).
5. "Sequential Bioluminescence Resonance Energy Transfer-Fluorescence Resonance Energy Transfer-Based Ratiometric Protease Assays with Fusion Proteins of Firefly Luciferase and Red Fluorescent Protein," B. R. Branchini, J. C. Rosenberg, D. M. Ablamsky, K. P. Taylor, T. L. Southworth and S. J. Linder*, *Analytical Biochemistry* 414: 239-245 (2011).
6. "Sensitive Dual-Color *In Vivo* Bioluminescence Imaging Using a New Red Codon Optimized Firefly Luciferase and a Green Click Beetle Luciferase," L Messanotte, E. Kaijzel, B. Branchini, A. Roda and C. Lowik, *PloS ONE* (www.plosone.org) 6 (4): e19277 (2011).
7. "Designing Quantum Rods for Optimized Energy Transfer with Firefly Luciferase Enzymes," R. Alam, D. Fontaine, B. Branchini and M. Maye, Nano Letters: 3251-3256 (2012).
8. "Novel Heterocyclic Analogs of Firefly Luciferin", C. C. Woodroofe, P. L. Meisenheimer, D. H. Klaubert, Y. Kovic, J. C. Rosenberg, C.E. Behney, T. L. Southworth and B. R. Branchini, Biochemistry 51: 9807-9813 (2012).

Patents:

1."Isolated Luciferase Gene of *L. italica*," Bruce R. Branchini, Tara L. Southworth, Jennifer P. DeAngelis, Aldo Roda and Elisa Michelini, US Patent number 7,807,429 B2, October 5, 2010.

2. "Chimeric Luciferases", Bruce R. Branchini, Provisional Patent filed, May 24, 2012.

Major Accomplishment 1:

The plasmid and technology for the expression, purification, nIR dye labeling and BRET activation of Ppy RE10 were transferred to Rajesh Naik's lab at AFRL.

Major Accomplishment 2:

The design, construction and evaluation of novel protease substrates for caspase and the blood clotting factors thrombin and factor Xa was carried out. The fusion protein substrates consist of a nIR-labeled red fluorescent protein linked to a thermostable luciferase through a peptide linker containing specific proteolytic cleavage sites. Signals are produced by a unique sequential BRET-FRET process and nanogram detection limits have been achieved. To progress toward making reagents that will function in blood by taking advantage of the superior transmission properties of nIR light, we will design, construct and evaluate nIR-dye labeled split luciferase systems.

Major Accomplishment 3:

A new approach to developing novel near Infrared (nIR) emitting materials using a bioluminescence resonance energy transfer (BRET) approach was undertaken in collaboration with Mathew Maye, Syracuse University. Our lab prepared versions of green (Ppy GR-TS) and red (Ppy RE9) light-emitting thermostable firefly luciferase variants containing an N-terminal spacer with 6xHis tag for covalent attachment to quantum rods fabricated in the Maye lab. The materials and technology for activating the bioluminescence sources were transferred to the Maye lab. The development of near Infrared (nIR) emitting materials using bioluminescence resonance energy transfer (BRET) from luciferases from our lab and quantum dots/rods from Mat Maye's laboratory at Syracuse University continues. This work has continued and has produced another paper in the journal *Nanoscale* ("Novel Multistep BRET-FRET Energy Transfer using Nanoconjugates of Firefly Proteins, Quantum Dots, and Red Fluorescent Proteins" Rabeka Alam, Joshua Zylstra, Danielle M. Fontaine, Bruce R. Branchini, Mathew M. Maye) that has been accepted for publication. For this work, the Maye lab used a new substrate analog that we supplied called benzothiophene luciferin ($BtLH_2$). The development of this new luciferase substrate is described under Major Accomplishment 4.

Major Accomplishment 4:

We evaluated new luciferin substrates with the goal of tailoring a substrate-enzyme combination to provide optimal BRET excitation sources. We will engineer a luciferase mutant with improved brightness and blue-shifted emission with a benzothiophene analog providing better spectral overlap. We are continuing a collaboration with the Maye lab and will explore improving BRET efficiency with firefly luciferin analogs.

OBJECTIVE 2. To develop and characterize new bioluminescent systems.

Publication:
"Chemical analysis of the luminous slime from the marine worm Chaetopterus (Annelida, Polychaeta)" Bruce R. Branchini, Curran E. Behney, Tara L. Southworth, Renu Rawat and Dimitri D. Deheyn, manuscript in preparation.

Major Accomplishment 5:

The marine annelid *Chaetopterus variopedatus* produces bioluminescence by an unknown and potentially novel mechanism. LC/MS and fluorescence analysis of extracts of harvested luminescent material revealed riboflavin as the major fluorescent component. A hypothesis is proposed in which an excited riboflavin derivative serves as the emitter in the worm's light producing reaction.

Two mycosporine-like amino acids with ~360 nm absorbance were isolated from the bioluminescent organs of *C. variopedatus* specimens provided by the Deheyn lab. LC/MS, UV-Vis and fluorescence spectroscopy suggested a possible a role in the bioluminescent process. However, these mycosporine-like amino acids were determined to be non-fluorescent and are therefore likely to be unrelated to light emission. A major ~20 kDa mucous protein band was isolated using SDS-PAGE. LC/MS, proteomics, and peptide sequencing revealed that the band contains ferritin as a minor (~2%) component and a major novel protein with mass 18, 427. LC/MS data were transferred to Deheyn lab for probe design in cloning approach.

Spectroscopic and LC/MS analyses of luminescent materials from *Chaetopterus variopedatus* provided by the Deheyn lab revealed the presence of riboflavin, a compound that fluoresces (λ_{max} = 535 nm). We determined that marine warm harbors *Vibrio campbellii* a bioluminescent bacterium that uses reduced flavin $FMNH_2$, a substrate made from riboflavin. However, this finding was not reproducible when fresh specimens were examined in a subsequent experiment. Evidently *V. campbellii* is a parasite that sometimes is found in *Chaetopterus*.

OBJECTIVE 3. To further investigate the domain alternation mechanism in P. pyralis firefly luciferase (Luc) bioluminescence using site-directed mutagenesis, time-resolved fluorescence spectroscopy and crystallography.

Publications:

9. "Bioluminescence Is Produced from a Trapped Firefly Luciferase Conformation Predicted by the Domain Alternation Mechanism," B. R. Branchini, J. C. Rosenberg, D. M. Fontaine, T. L. Southworth, C. E. Behney, and L. Uzasci, *Journal of the American Chemical Society* 133: 11088-11091 (2011).
10. "Crystal Structure of Firefly Luciferase in a Second Catalytic Conformation Supports a Domain Alternation Mechanism," J. A. Sundlov, D. M. Fontaine, T. L. Southworth, B. R. Branchini, and A. M. Gulick, Biochemistry 51: 6493–6495 (2012).

Major Accomplishment 6:

According to the domain alternation mechanism and crystal structure evidence, the acyl-CoA synthetases, one of 3 subgroups of a superfamily of adenylating enzymes, catalyze adenylate- and thioester-forming partial reactions in two different conformations. The enzymes accomplish this by presenting two active sites through an ~140° rotation of the C-domain. The 2nd partial reaction catalyzed by another subgroup, the beetle luciferases, is a mechanistically dissimilar oxidative process that produces bioluminescence. We have demonstrated that a firefly luciferase variant containing cysteine residues at positions 108 and 447 can be intramolecularly cross-linked by 1,2-bismaleimidoethane, trapping the enzyme in a C-domain rotated conformation previously undocumented in the available luciferase crystal structures. The cross-linked luciferase cannot adenylate luciferin, but is nearly fully capable of bioluminescence with synthetic luciferyl-adenylate because it has retained the ability to carry out the oxidative half-reaction. The cross-linked luciferase has apparently been trapped in a conformation similar to those adopted by acyl-CoA synthetases as they convert acyl-adenylates into the corresponding CoA thioesters.

Major Accomplishment 7:

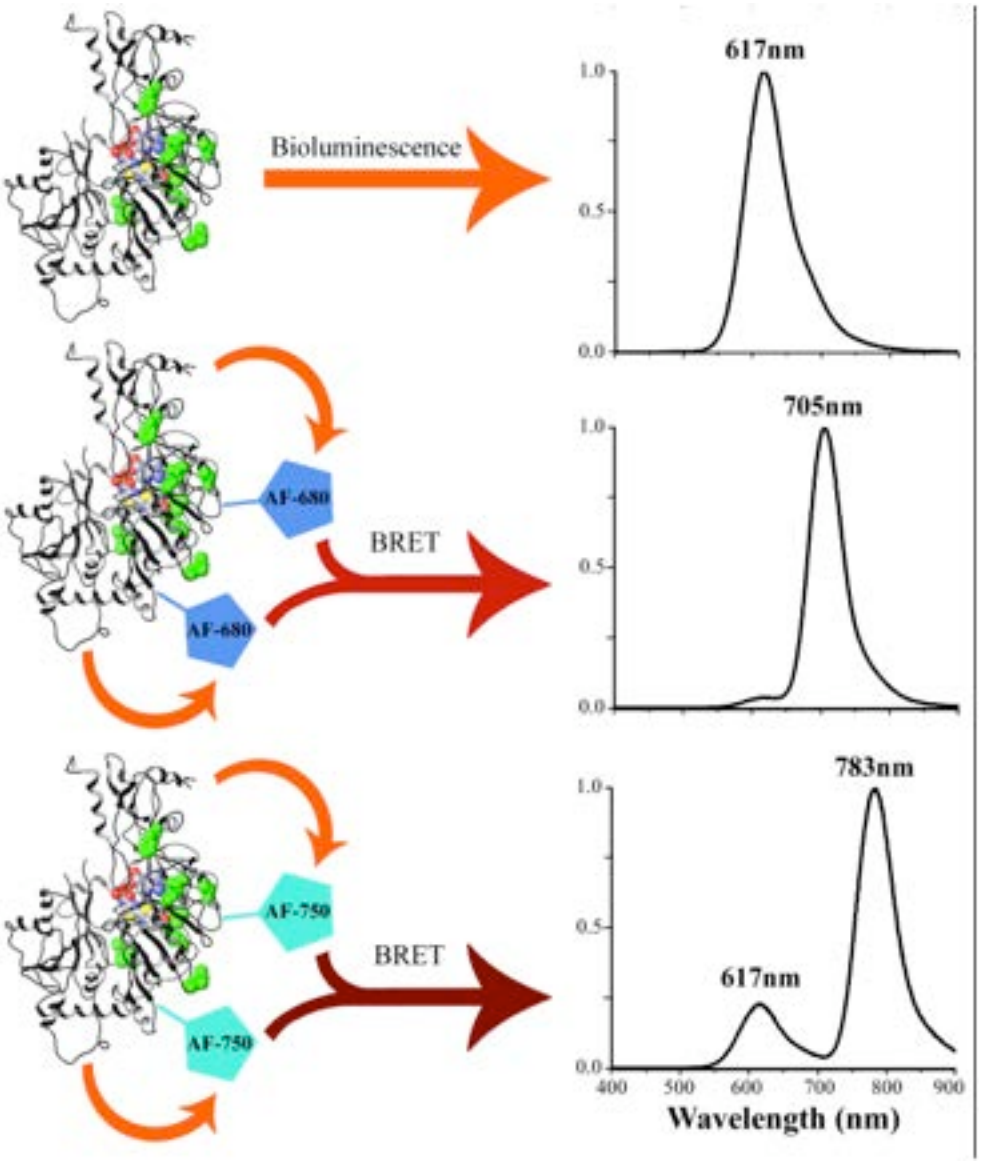
In collaboration with A. Gulick (Hauptman Woodward Institute) we have determined the crystal structure of a cross-linked luciferase that is trapped in a conformation that catalyzes the oxidative light emitting half-reaction. The new x-ray structure supports the role of the 2nd catalytic conformation and provides insights into the biochemical mechanism. These results provided very strong substantiating structural evidence for the domain alternation mechanism of firefly bioluminescence. Mechanistic investigations stimulated by the new crystal structure are underway.

III. Detailed Descriptions of the 7 Major Accomplishments.

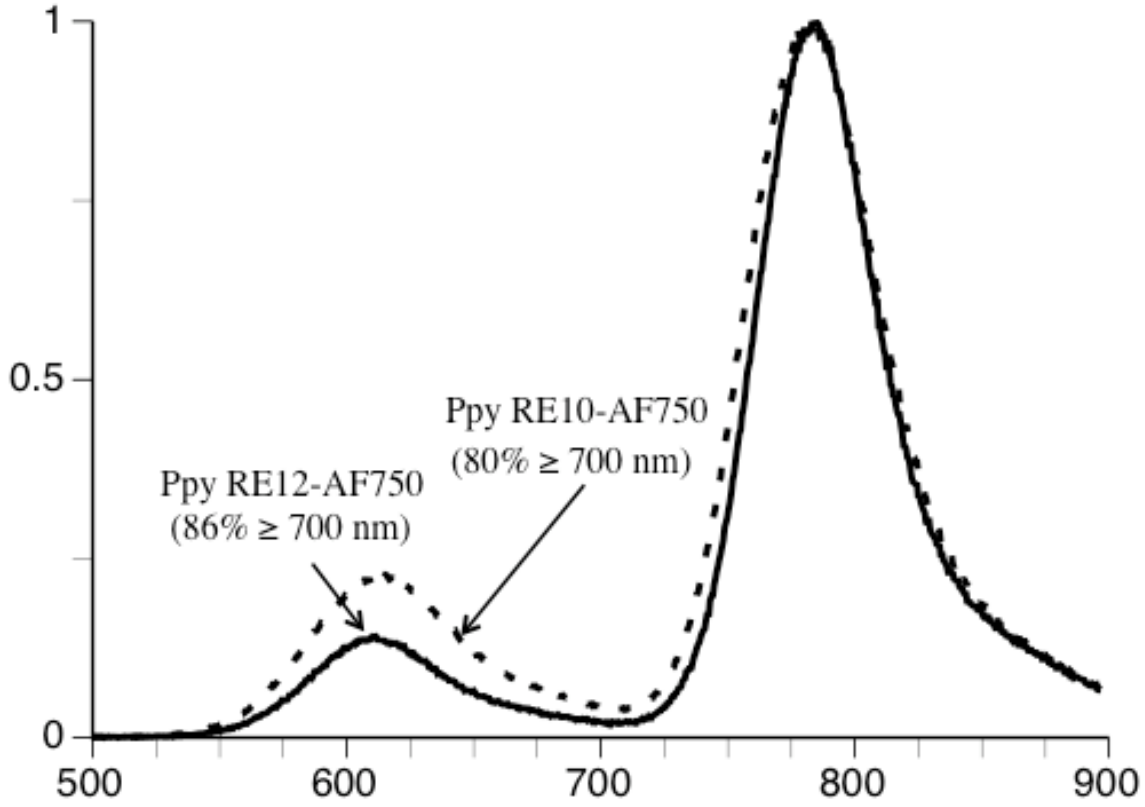
Major Accomplishment 1:

The plasmid and technology for the expression, purification, nIR dye labeling and BRET activation of Ppy RE10 were transferred to Rajesh Naik's lab at AFRL, Dayton, Ohio.

Progress has been made toward improving the BRET efficiency (to minimize residual visible light emission) of a long wavelength nIR source made by AF-750 labeling. Previously, we had prepared a luciferase variant Ppy RE10 in which 2 surface Cys residues were chemically labeled with nearer (nIR) fluorescent dyes AF-680 and AF-750. The materials produced efficient nIR output through the process of bioluminescence resonance energy transfer (BRET) as shown in the figure below.



We had some success in improving the efficiency of the BRET transfer in the desirable longer wavelength emitter, although our goal of eliminating visible emission was not met. The approach we used was to engineer two additional surface Cys residues into the Ppy RE10 mutant yielding a new luciferase variant Ppy RE12. After labeling with AF750, a 30% decrease in the residual visible bioluminescence (compared to the results with Ppy RE 10) was achieved as shown in the diagram below.



We believe that the practical limitations of the multiple dye labeling (solubility of the protein in buffer decreases) and the sequential BRET-FRET approaches to minimizing the visible light production from nIR-emitting luciferases have been reached. These efforts will not be continued. The future progress (with direct BRET) will require new nIR fluorescent dye maleimides with emission maxima > 750 nm and higher quantum yields of at least ~ 0.3.

Major Accomplishment 2.

A diverse group of biological organisms are bioluminescent, emitting light that spans the visible spectrum (*1, 2*). The ability to engineer bioluminescent proteins has expanded the already impressive number of analytical applications that are possible with these fascinating, glowing materials (*3, 4*). Our interest is in developing new biomaterials for improved biomarker and biosensor applications, focusing on the firefly luciferase system from *Photinus pyralis* (Luc[1]). The beetle enzyme produces yellow-green light (λ_{max} = 560 nm) through a series of reactions that require substrates firefly luciferin (LH_2), Mg-ATP and oxygen. This process is accomplished with a high quantum yield of 41 ± 7.4% (*5*), expressed as the conversion of substrate LH_2 into photons. An additional advantage of working with Luc is the excellent stability and relatively low cost of LH_2.

Recently, we reported (*6*) on the design, construction and characterization of chemically modified Luc variants that emitted near-infrared (nIR) light via the bioluminescence resonance energy transfer (BRET) process. BRET and fluorescence resonance energy transfer (FRET) involve non-radiative energy transfer from a bioluminescent or fluorescent donor to an acceptor. The efficiency of these processes depends on the spectral overlap, relative orientation and inverse sixth-power of the distance between the donor and acceptor (*7*). For BRET, efficiency can be expressed as the ratio of fluorescence acceptor emission to bioluminescence emission (*8*). In one example (*6*), we constructed BXRE-680, a fusion protein consisting of a biotin-binding domain joined through a decapeptide linker to a Luc variant covalently labeled with the nIR fluorescent dye Alexa Fluor 680 C2-maleimide (AF680). Upon addition of LH_2 and Mg-ATP, the fusion protein produced an emission spectrum with a maximum at 706 nm as a result of an efficient intramolecular BRET process (BRET ratio = 14.0). The very favorable BRET efficiency of BXRE-680 was achieved by: (1) the good overlap between the red bioluminescence emission (λ_{max} = 617 nm) of the Luc variant and the AF680 absorbance (λ_{max} = 680 nm); and (2) the attachment of the nIR dyes to Cys residues ~20 Å and ~35 Å from the emitter site. When immobilized onto agarose-streptavidin, BXRE-680 was used to assay physiological concentrations (1-20 μg/ml) of factor Xa (*6*). The basis of the assay was that cleavage of BXRE-680 at the factor Xa specific IleGluGlyArg site in the linker released the AF680-labeled Luc variant into solution. The high transmittance of the nIR luminescence through blood enabled model assays of 1 μg/ml levels of factor Xa in this medium.

Building on our prior work with Luc sources of nIR (*6*), we sought to develop luminescent reagents for enzyme assays that could be used in a ratiometric format taking advantage of the high sensitivity and low background of bioluminescence, while avoiding difficulties with the standardization of intensity units. Ratiometric probes require two distinct signals and are self-referencing because the ratio of the two signals is independent of the probe concentration (*9, 10*). Recent reports have highlighted ratiometric methods for detecting DNA and micro-RNA (*11*), strategies for the design of nIR ratiometric fluorescent probes (*12*), ratiometric FRET measurements in living cells using microscopic imaging systems (*13*), and ratiometric bioluminescence–based indicators for Ca^{2+} imaging (*14*) and cAMP monitoring *in vivo* (*15*). Our approach was to create ratiometric luminescent probes for monitoring proteolytic activity that contain two well-separated emission peaks produced by a sequential BRET-FRET process. To the best of our knowledge, our application is novel and is the first sequential BRET-FRET process contained in a single protein. We note, however, that the first sequential BRET-FRET process, termed SRET (*16*), involved the detection of ternary complex formation by three

different interacting fusion proteins through signals produced by *Renilla* luciferase and two fluorescent protein acceptors, each fused to one of the proteins of interest. Our probes are single soluble fusion proteins consisting of a red fluorescent protein (RFP) labeled with AF680 and Ppy WT-TS (*17*), a thermostable Luc variant that catalyzes 560 nm bioluminescence. The two proteins are connected by a decapeptide containing a protease recognition site (Fig. 1). We explored the potential utility of the assay concept with representative serine proteases factor Xa and thrombin involved in blood coagulation and the cysteine protease caspase 3, which plays a key role in apoptosis (*18*). BRET- or FRET-based reagents for monitoring the activity of thrombin and caspase 3 have been reported (*19-22*); and the proteases are relevant to monitoring heparin anticoagulant treatment (*23*), therapeutic agent development and cancer research (*24*). We present here studies on the design and construction of three specific protease substrates, the intramolecular SRET-based assay principle and results that indicate excellent sensitivity and selectivity can be achieved with this methodology in a ratiometric format.

Materials and methods

Materials

The following materials were obtained from the sources indicated: Mg-ATP (bacterial source), human thrombin and human caspase 3 from Sigma-Aldrich (St. Louis, MO); restriction endonucleases, TPCK-treated trypsin and factor Xa protease (bovine) from New England Biolabs (Beverly, MA); Alexa Fluor® 680 C2-maleimide (AF680) dye from Invitrogen (Carlsbad, CA). LH_2 was a generous gift from Promega (Madison, WI). The recombinant GST-fusion proteins Ppy WT and Ppy WT-TS were expressed and purified as previously reported (*17*). The mKate S158A variant gene (GenBank Accession # EU383029) flanked by *Bam*HI and *Xho*I restriction endonuclease sites in the pUC57 vector was synthesized by GenScript USA Inc (Piscataway, NJ). The found molecular masses (Da) of the following proteins were within the allowable experimental error (0.01%) of the calculated values (in parenthesis): BFFP-Xa, 89 131 (89 138); BFFP-C3, 89 138 (89 141); BFFP-Th, 89 142 (89 148); BFFP-GS, 88 962 (88 971); RFP, 26 596 (26 599); RFP C114S, 26 581 (26 583); RFP C222S, 26 582 (26 583); and RFP C114S/C222S, 26 564 (26 567).

General methods

Concentrations of Ppy WT and Ppy WT-TS were determined with the Bio-Rad Protein Assay system using BSA as the standard. Specific activity measurements were determined as previously reported (*25, 26*) except that the integration time was 15 min and the final LH_2 concentration was 0.5 mM. Fusion protein substrate concentrations were determined by UV using the RFP chromophore ($\epsilon_{585\ nm}$ = 51 000 M^{-1} cm^{-1} for the mKate S158A variant (*27*)) after the $A_{585\ nm}$ was reduced by 19% of the maximum absorbance of AF680 to correct for spectral overlap. Mass spectral analyses were performed by tandem HPLC-electrospray ionization mass spectrometry (LC/ESMS) using a ThermoFinnigan Surveyor HPLC system and a ThermoFinnigan LCQ Advantage mass spectrometer. Post run data analysis was performed using ThermoFinnigan BioWorks Browser 3.0 deconvolution software.

Covalent labeling with AF680 to produce fusion protein substrates

Stock solutions of AF680 (10 mM, determined by UV using ϵ_{684nm} = 175 000 M^{-1} cm^{-1} (*28*)) were prepared in sterile deionized water, divided into 30 µl aliquots, lyophilized and stored

at -20 °C. The labeling reactions to produce BFS-C3, BFS-Th, BFS-Xa and BFS-GS were performed at 10 °C in 20 mM sodium phosphate buffer (pH 7.2) containing 150 mM NaCl, 5 mM EDTA and 0.8 M ammonium sulfate (PBSA buffer). Ammonium sulfate was included because it substantially reduced non-covalent dye incorporation of nIR dye. Labeling reactions were initiated by addition of 1 ml of 30 μM fusion protein in PBSA buffer to a lyophilized aliquot of AF680 (0.3 mM final concentration), gently mixed, and incubated for 30 min. Reactions were quenched by addition of 20 μl of glutathione solution (2 mM final concentration), and after 15 min were exhaustively dialyzed against PBSA buffer (8 changes, 1 L each). The dye-labeled fusion proteins could be stored in PBSA buffer at 5 °C for at least 2 months without loss of more than 10% activity.

The dye/protein ratios of the fusion protein substrates were estimated from the μmol of AF680 determined by UV using ϵ_{684nm} = 175 000 M^{-1} cm^{-1} (*28*) and the μmol of protein determined by UV using $\epsilon_{585\ nm}$ = 51 000 M^{-1} cm^{-1} for the mKate S158A variant (*27*). Corrections were applied to the absorbance readings to compensate for spectral overlap. The substrate nIR dye/protein ratios were 1.8 ± 0.3. Each fusion protein substrate was also analyzed by LC/ESMS using a BioBasic-C4 (100 x 1 mm) column eluted at a flow rate of 50 μl/min with an acetonitrile gradient of 10%/min. Absorbance was monitored at 280 nm and 680 nm and the long wavelength signal was only observed co-eluting with protein indicating the absence of any non-covalently incorporated AF680. The MS analysis of each substrate produced two signals in ~ 2:1 ratio corresponding to mass increases of 1958 Da and 979 Da, respectively. Based on the mass of AF680 (979 Da), the mass increases of 1958 Da and 979 Da corresponded to 2:1 and 1:1 dye/protein incorporation ratios, respectively.

Luminescence spectroscopy

Emission spectra produced by bioluminescence, BRET and SRET were obtained using a Horiba Jobin-Yvon *i*HR imaging spectrometer equipped with a liquid N_2 cooled CCD detector and the excitation source turned off. Data were collected at 25 ºC in a 0.8 ml quartz cuvette over the wavelength range 400-935 nm with the emission slit width set to 25 nm. Fluorescence emission spectra were measured with the excitation source set to 585 nm and the excitation and emission slits set to 1 nm. All spectra were corrected for the spectral response of the CCD using a correction curve provided by the manufacturer. BRET ratios of the fusion proteins prior to labeling with AF680 were estimated by dividing the BRET peak intensity at 610 nm by the residual bioluminescence peak intensity at 565 nm. FRET efficiencies were determined by comparing fluorescence spectra of 1.0 μM solutions of fusion proteins prior to labeling with AF680 with the corresponding fusion protein substrate (labeled with AF680). The FRET efficiencies of the fluorescence at 625 nm from RFP in the unlabeled fusion proteins (F_D) and in the corresponding substrates ($F_{D\text{-}A}$) were calculated according to the equation $E = 1 - F_{D\text{-}A}/F_D$ (*7*). SRET ratios for the fusion proteins were calculated from emission spectra by dividing the intensity of the residual bioluminescence peak at 560 nm by the intensity of the SRET emission peak at 705 nm.

Proteolytic cleavage of BFS-Xa

The specific activity of the commercial factor Xa protease sample first was confirmed by an absorbance assay using the chromogenic substrate Bz-Ile-Glu(γ-OR)-Gly-Arg-pNA·HCl (Kabi Diagnostica, Stockholm, Sweden) performed according to the manufacturer's directions. The time course of cleavage of BFS-Xa (2.0 μM) by factor Xa (0.2 μM) at 20 °C in 0.05 ml

PBSA was monitored by withdrawing 5 μl aliquots of incubation mixtures and adding them to 0.50 ml of 25 mM glycylglycine buffered (pH 7.8) solutions of LH_2 (150 μM) and Mg-ATP (2 mM). After mixing for 10 s, emission spectra were recorded (see above) and are shown in Fig. 4.

Substrate specificity

The cleavage of BFS-C3 by caspase 3, BFS-Th by thrombin and BFS-Xa by factor Xa at the expected sites shown in Fig. 1 was confirmed by analyzing the products resulting from exhaustive proteolytic digestion of each substrate. Proteases (1.0 μM) were added to 0.1 ml volumes of ~20 μM fusion protein substrates in PBSA buffer until emission spectra indicated the complete disappearance of the long wavelength SRET peak. LC/ESMS analysis of the digested fusion protein solutions was performed using a BioBasic-C4 (100 x 1 mm) column eluted at a flow rate of 50 μl/min using an acetonitrile gradient of 10%/min.

The relative rates of cleavage of the fusion protein substrates were determined by incubating each one (2.0 μM) at 20 °C with 2.0 μM protease (caspase 3, thrombin and factor Xa) in PBSA buffer. Aliquots (5 μl) were withdrawn at various intervals over 30 min and added to cuvettes containing 0.50 ml of a 25 mM glycylglycine buffered (pH 7.8) solution of LH_2 (150 μM) and Mg-ATP (2 mM). Samples were briefly mixed and placed in the sample compartment of a Horiba Jobin-Yvon *i*HR imaging spectrometer and emission spectra were recorded as detailed above. The changes in 560 nm/760 nm peak ratios per min were determined and are expressed as % of the rate obtained with BFS-C3 and caspase 3 (0.58/min), defined as 100%.

Ratiometric assays of caspase 3, factor Xa and thrombin activity

The rates of protease cleavage of fusion protein substrates were monitored by changes in luminescence intensity ratios. Mixtures (0.05 ml) of 2.0 μM fusion protein substrates and proteases (BFS-C3 with 0.408 nM to 8.17 nM caspase 3; BFS-Xa with 58.0 nM to 465.0 nM factor Xa; and BFS-Th with 1.0 nM to 20.0 nM thrombin) were incubated in PBSA at 20 °C and 5 μl aliquots were withdrawn at the times indicated in Fig. 5 and added to cuvettes containing 0.50 ml of a 25 mM glycylglycine buffered (pH 7.8) solution of LH_2 (150 μM) and Mg-ATP (2 mM). Samples were briefly mixed and placed in the sample compartment of a Horiba Jobin-Yvon *i*HR imaging spectrometer and emission spectra were recorded. Peak ratios were determined from the intensities of the 560 nm (bioluminescence) and 706 nm (SRET) peaks.

The relationships between protease concentration and change in luminescence intensity ratio were determined using both the imaging spectrometer and a PerkinElmer LS55 luminescence spectrometer in endpoint analysis assay format. Mixtures (0.05 ml) of fusion protein substrates and varying concentrations (Fig. 6) of target proteases were incubated in PBSA at 20 °C for 15 min and 5 or 10 μl aliquots were withdrawn and added to cuvettes containing 0.50 ml of a 25 mM glycylglycine buffered (pH 7.8) solution of LH_2 (150 μM) and Mg-ATP (2 mM). Samples were briefly mixed and placed in the sample compartment of a Horiba Jobin-Yvon *i*HR imaging spectrometer or after 1 min (to ensure slow decay kinetics as shown in Fig. 3) into the scanning fluorometer and emission spectra were recorded as detailed above. The Perkin Elmer spectrometer was operated in the "bioluminescence" mode as previously described (*29*) using a wavelength range of 500 nm to 750 nm and a scan speed of 250 nm/min. The excitation source was turned off and the emission slit was set to 10 nm.

Cloning of the fusion proteins BFFP-Xa, BFFP-C3, BFFP-Th and BFFP-GS

The vector containing the gene encoding the fusion protein BFFP-Xa, consisting of an N-terminus hex-His tagged mKate S158A variant joined to Ppy WT-TS through the decapeptide linker GSAIEGRGSG, was constructed as follows. The synthesized mKate S158A variant gene (above) in the pUC57 vector was digested with *Bam*HI and *Xho*I and ligated into the corresponding cloning sites in a modified version of the pQE-30 plasmid (Qiagen, Valencia, CA). A Gly codon and *Afe*I site were then inserted prior to the stop codon of the mKate gene using the following primer and its respective reverse complement: 5´- A CTG GGC CAT AAA CTG AAC **GGA AGC GCT** TAA CTC GAG ACC CCG GGT CG -3´ (bold represents the inserted codons and underline represents the *Afe*I site). The Ppy WT-TS gene in the pGEX-6P-2 expression vector was modified by inserting an *Afe*I restriction site followed by the codons for linker residues IEGRGSG upstream of the start codon using the following primer and its respective reverse complement: 5´– G CCC CTG GGA TCC **AGC GCT ATC GAA GGT CGT GGA TCC GGA** ATG GAA GAC GCC AAA AAC – 3´ (bold represents the inserted codons, underlined represents the *Afe*I site). The plasmids containing the altered mKate and Ppy WT-TS genes were digested with *Afe*I and *Xho*I and, after purification with a Qiagen QIAquick kit, the modified Ppy WT-TS gene fragment was ligated into the pQE-30 plasmid producing the expression vector for BFFP-Xa. The vectors containing the genes encoding the fusion proteins BFFP-C3, BFFP-Th and BFFP-GS were prepared from the corresponding BFFP-Xa expression vector by mutating the codons for the tetrapeptide protease recognition site within the decapeptide linker with the QuikChange® Lightning Site-Directed Mutagenesis kit. The following primers and their respective reverse compliments were used to make: BFFP-C3 by introducing DEVD, 5´- CTG AAC GGA AGC **GCA GAC GAA GTT GAT** GGA TCC GGA ATG GAA GAC GCC -3´; BFFP-Th by introducing LVPR, 5´- CTG AAC GGA AGC **GCA CTC GTA CCT CGT** GGA TCC GGA ATG GAA GAC GCC -3´; and BFFP-GS by introducing GSGS, 5´- CTG AAC GGA AGC **GCA GGC TCA GGT AGT** GGA TCC GGA ATG GAA GAC GCC AAA AAC -3´(where bold represents the mutated bases, and underline represents the removal of the *Afe*I site).

Identification of RFP residues labeled with nIR dye

The labeling sites of the fusion protein substrates with AF680 were indirectly identified as Cys126 and Cys234 as indicated in Fig. 1 in the main text. These residues correspond to positions 114 and 222 of RFP. Covalent labeling of RFP with nIR dye (for 60 min at 10 °C) and LC/ESMS analysis, performed as described in the main text, revealed a major product with a mass increase of 1958 Da consistent with the covalent attachment of 2 mol of nIR dye/ mol RFP. Based on molecular modeling analysis of an RFP crystal structure, Cys114 and Cys222 appeared to be surface accessible, while the thiols of Cys 31 and Cys172 seemed to be oriented toward the center of the β-barrel structure. Three RFP variants C114S, C222S and C114S/C222S were made by site directed mutagenesis and the proteins were expressed and purified (see above). After labeling with nIR dye and subsequent LC/ESMS analysis, the single Cys mutants had mass increases of 979, while the double mutant mass was unchanged, failing to incorporate AF680. These results are consistent with Cys 114 and Cys222 being the primary sites of AF680 labeling. Presumably, the equivalent residues in the fusion protein substrates (Cys 126 and Cys234) are similarly labeled. Based on an analysis of the masses of the fusion protein substrate proteolysis products (see Table 2 in main text), the Ppy WT-TS portion of the substrates was not labeled

with AF680. This result is consistent with previously reported control experiments that indicated the four native Cys residues of Luc are not accessible to AF680 under labeling conditions similar to those employed in this study.

Results and discussion

Design and initial characterization of SRET protease substrates

To demonstrate the applicability of the SRET principle to monitoring enzyme activity, we prepared the soluble fusion protein substrates shown in Fig. 1 that contain covalently attached nIR dye AF680. The cloning, expression and purification of the precursor His-tagged fusion proteins in yields of 3.5 to 4.5 mg/0.25 L culture are described elsewhere. Prior to labeling with AF680, addition of LH_2 and Mg-ATP to the proteins produced overlapping emission peaks of approximately equal intensity with maxima at 565 nm and 610 nm (Fig. 2c)

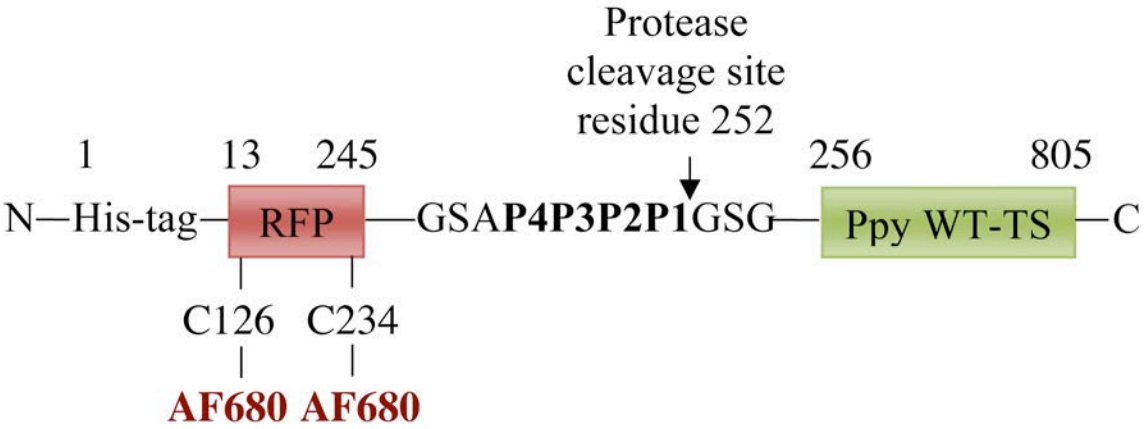


Fig. 1. Schematic representation of the fusion protein substrates. P4P3P2P1 corresponds to: BFS-Xa, IEGR; BFS-C3, DEVD; BFS-Th, LVPR and BFS-GS, GSGS. The arrow indicates the cleavage site for the three proteases. BFS-GS, which does not contain a specific protease site, served as a control for substrate specificity and stability. The numbers above the schematic diagram indicate the amino acid sequence of the fusion proteins.

representing residual bioluminescence and BRET, respectively. The overlap of the Ppy WT-TS donor emission spectrum (Fig. 2a) and the absorption spectrum of the acceptor RFP (Fig. 2b) is very good; so it is likely that the modest BRET efficiencies (Table 1) are mainly the result of the distance between the two interacting species. While the length of the peptide linking the RFP to Ppy WT-TS was not optimized, it is unlikely that it can be shortened because it contains only tripeptide spacers flanking the 4 amino acids comprising the protease sites that must be readily accessible to the proteases being monitored.

Table 1
Specific activities and spectral properties of fusion protein substrates.

Protein	Relative integrated specific activity [a] (15 min)	Emission maxima (nm ± 1)	Peak ratio (560 nm/706 nm)	BRET ratio [c]	FRET efficiency [d]
Ppy WT	100	560	-	-	-
Ppy WT-TS	102	560	-	-	-
BFS-Xa	103	562, 707 [b]	0.70	0.92	0.83
BFS-Th	97	562, 705 [b]	0.88	0.82	0.82
BFS-C3	93	562, 705 [b]	0.73	0.82	0.86
BFS-GS	95	560, 706 [b]	0.78	0.84	0.83

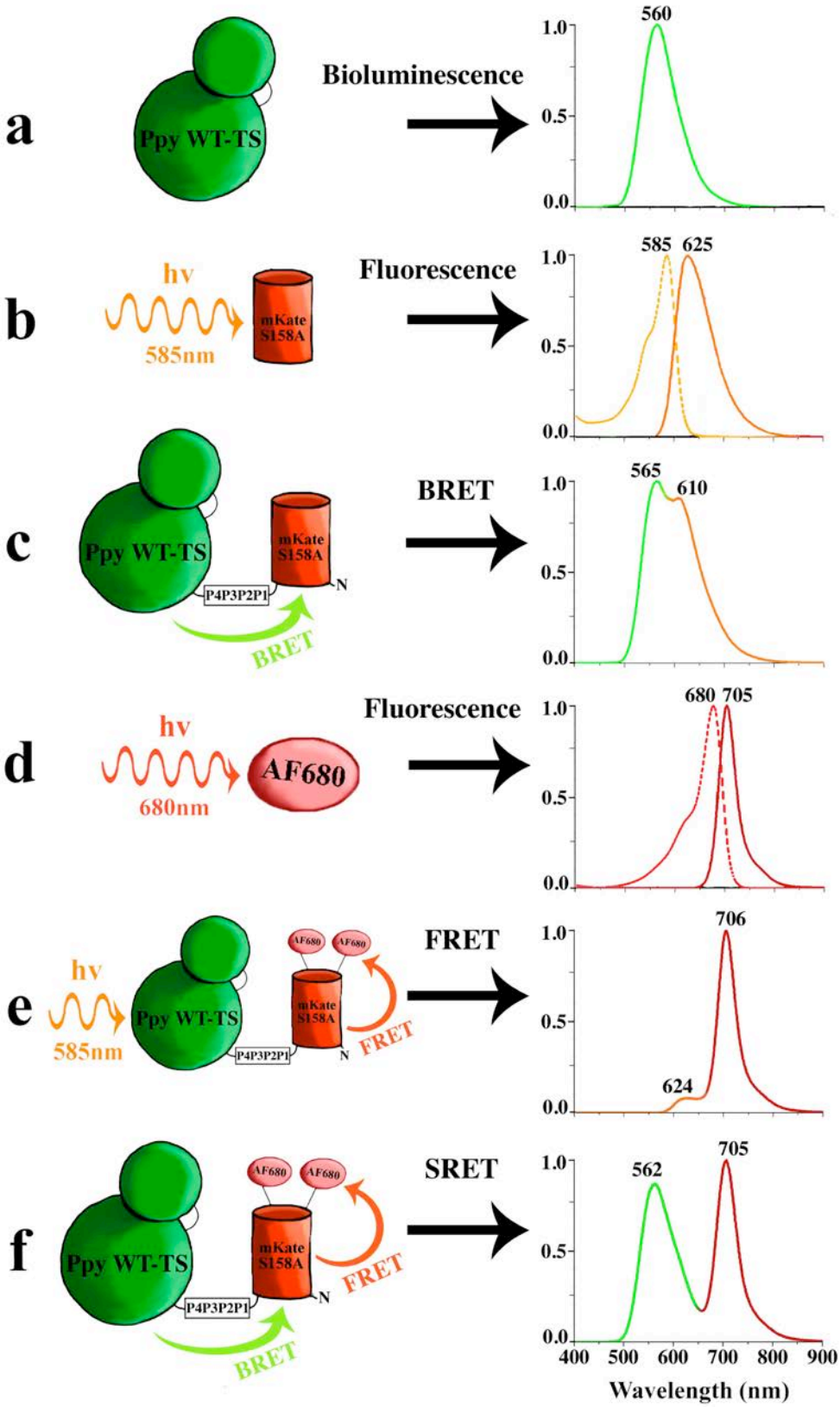

a
Ppy WT-TS
Bioluminescence
560
1.0
0.5
0.0
b
hv
585nm
mKate S158A
Fluorescence
585
625
1.0
0.5
0.0
c
Ppy WT-TS
mKate S158A
P4P3P2P1
N
BRET
BRET
565
610
1.0
0.5
0.0
d
hv
680nm
AF680
Fluorescence
680
705
1.0
0.5
0.0
e
hv
585nm
Ppy WT-TS
AF680
AF680
mKate S158A
P4P3P2P1
N
FRET
FRET
706
624
1.0
0.5
0.0
f
Ppy WT-TS
AF680
AF680
mKate S158A
P4P3P2P1
N
FRET
BRET
SRET
562
705
1.0
0.5
0.0
400 500 600 700 800 900
Wavelength (nm)

Fig. 2. Normalized absorption, emission and fluorescence spectra. The following spectra were obtained as described in Materials and methods and demonstrate: (a) the bioluminescence of Ppy WT-TS following addition of LH_2 and Mg-ATP; (b) RFP absorbance and fluorescence (excitation 585 nm); (c) BRET produced by unlabeled fusion protein substrate following addition of LH_2 and Mg-ATP; (d) AF680 absorbance and fluorescence (excitation 680 nm); (e) FRET produced by 585 nm excitation of fusion protein substrate; (f) SRET of BFS-Th following addition of LH_2 and Mg-ATP.

Table 2
Analysis of fusion protein substrate proteolysis products.

Fusion Protein	Protease	Found mass (Da) [a]	Identity of peptide	Calculated mass (Da)
BFS-Xa	Factor Xa	29,251 ± 3, 30,230 ± 3 60,876 ± 6	Met1-Arg252 [b] Gly253-Leu805	29,254, 30,233 60,881
BFS-C3	Caspase 3	29,254 ± 3, 30,233 ± 3 60,879 ± 6	Met1-Asp252 [b] Gly253-Leu805	29,257, 30,236 60,881
BFS-Th	Thrombin	29,262 ± 3, 30,240 ± 3 60,880 ± 6	Met1-Arg252 [b] Gly253-Leu805	29,264, 30,243 60,881
BFS-GS	Trypsin [c]	27,635 ± 3, 28,615 ± 3 49,732 ± 5	Gly3-Lys239 [b] Leu244-Lys694	27,633, 28,612 49, 734

[a] Determined from LC/ESMS analysis of exhaustive proteolytic digestion as described in Materials and methods.
[b] Found and calculated masses correspond to indicated peptide sequence containing one or two attached AF680 labels (979 Da).
[c] Performed under limited proteolysis conditions.

Ratiometric probes, which have two distinct signals in the presence or absence of analyte, offer many potential advantages in enzyme assays (*9*). Among them is the elimination of the need to accurately measure substrate concentration or to work at saturating levels of probe. Moreover, for luminescence-based methods, it is unnecessary to make absolute light intensity measurements avoiding problems associated with instrument standardization that make it difficult to compare results from different laboratories. Since we intended to develop ratiometric assays, we explored the possibility that attachment of AF680 to RFP would produce a FRET signal that would shift the BRET peak to longer wavelength, thereby improving the separation of the bioluminescence and BRET peaks (Fig. 2c). To achieve satisfactory peak separation, RFP would have to be selectively labeled and the FRET transfer would have to be very efficient. The fusion proteins contain 8 free Cys thiols, 4 each from Ppy WT-TS and RFP. We had previously determined (*6*) that the 4 native Cys residues of Luc were not accessible to AF680 and this was confirmed by analysis of the nIR-labeled fusion protein substrates (Table 2). Instead, RFP was mainly doubly labeled at Cys126 and Cys234 (corresponding to positions 114 and 222 in the RFP sequence), although ~30% of the species present were singly labeled (either Cys126 or Cys234, Table). The labeling results were reproducible and no attempts were made to optimize them further. A control experiment in which RFP was treated with AF680 revealed that highly efficient FRET could, in fact, be realized (Fig. 2e). Results similar to those shown in Fig. 2e were obtained upon 585 nm irradiation of the fusion protein substrates (Table 1). The highly efficient FRET may be

attributed to: (1) the very good overlap of the donor fluorescence of RFP (Fig. 2b) and the absorption spectrum of the acceptor AF680 (Fig. 2d); and (2) the short transfer distances provided by the direct attachment of the nIR dye to the protein. The highly efficient intramolecular FRET process was key to the production of fusion protein substrates that emit two well-separated peaks having maxima of approximately 560 nm and 705 nm (Fig. 2f), with the latter emission arising from the SRET process.

The four fusion protein substrates (Fig. 1) vary only in the tetrapeptide protease site and are designated as follows: BFS-Xa (IEGR), BFS-C3 (DEVD), BFS-Th (LVPR) and BFS-GS (GSGS). The tetrapeptides IEGR, DEVD and LVPR provided consensus sites for specific cleavage by factor Xa, caspase 3 and thrombin, respectively. The GSGS sequence was included to provide a control for substrate specificity and stability.

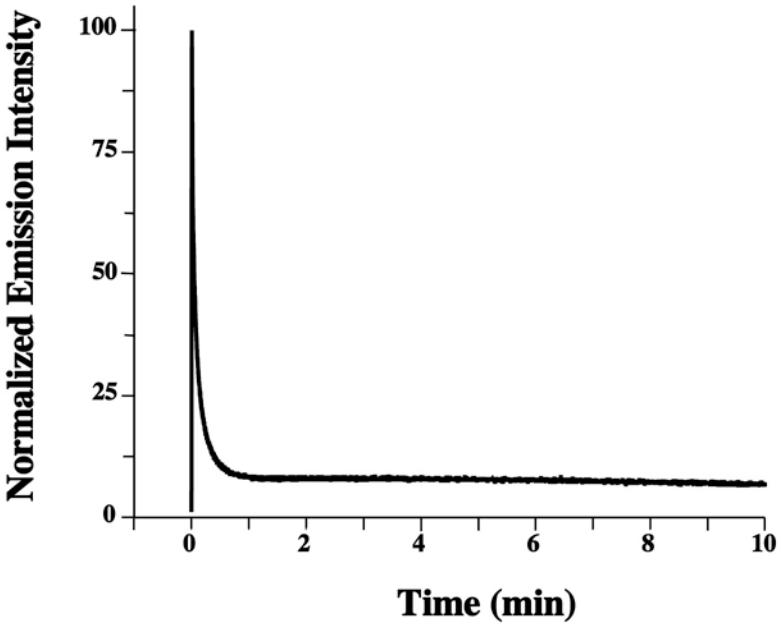


Fig. 3. Light emission profile of BFS-Xa. The intensity of the light emission produced by the fusion protein substrates, illustrated here with BFS-Xa, was monitored with a custom-built luminometer containing a Hamamatsu R928 photomultiplier tube (*31*). The reaction was initiated by the injection of 0.12 ml of 9.0 mM Mg-ATP into 0.40 ml of a 25 mM glycylglycine buffered mixture (pH 7.8) of LH_2 (0.7 μM) and 0.46 μg BFS-Xa.

The luminescence-based specific activities of the fusion protein substrates were quite similar to each other, nearly maintaining the exceedingly high activity of Ppy WT (Table 1). Excellent FRET efficiencies, modest BRET ratios and similar SRET maxima with minor variations in peak ratios characterized these novel materials (Table 1). The light emission kinetics of all the substrates, illustrated by BFS-Xa in Fig. 3, were identical and are characterized by a rapid flash followed by flat response after 1 min. Although the steady signal provides sufficient sensitivity that can readily be advantageously monitored with a recording fluorometer, the kinetic profile could be further optimized to provide enhanced sensitivity.

An alternative substrate design was also explored in which the luciferase and RFP proteins were placed at the N- and C-terminal domains of the fusion protein, respectively. After

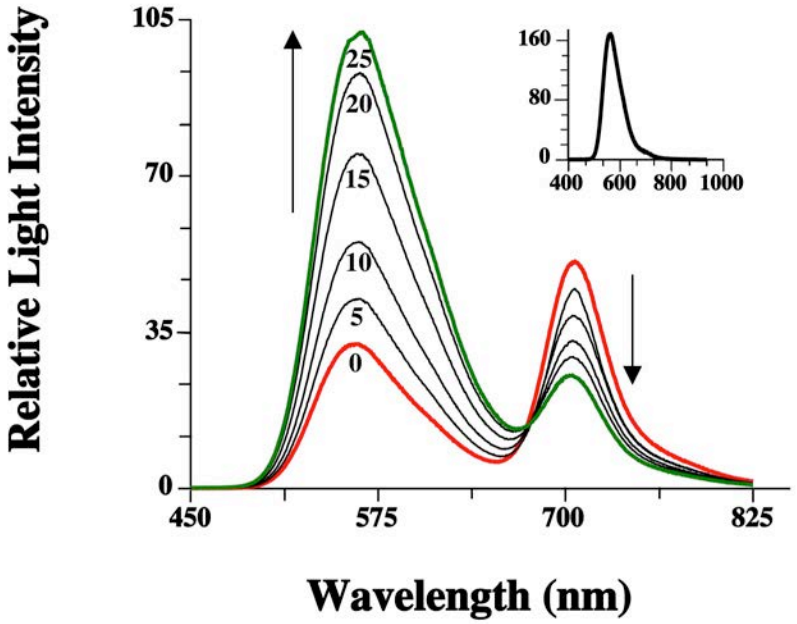


Fig. 4. Time course of cleavage of BFS-Xa by factor Xa. BFS-Xa (2.0 μM) and factor Xa (0.2 μM) were incubated at 20 °C in 0.05 ml PBSA. At the times (min) indicated in the figure, emission spectra were initiated by adding 5 μl aliquots of reaction mixtures to glycylglycine buffered (pH 7.8) solutions of LH_2 and Mg-ATP as described in detail in Materials and methods. The inset shows the emission spectrum obtained from a 5 μl aliquot of the cleavage reaction mixture after 24 h of incubation at 20 °C.

labeling RFP with AF680, a fusion protein substrate was made that was essentially BFS-Xa with the RFP and Ppy WT-TS sequences swapped. The substrate had very similar luminescent properties compared to BFS-Xa, however, because the rate of cleavage by factor Xa was somewhat slower with the linker peptide attached to the C-terminus of Ppy WT-TS (data not shown), we abandoned this approach. The similar luminescent properties of the N- and C-terminal attached RFP fusion proteins is consistent with the report (*30*) that fusion proteins of *Pyrocoelia miyako* luciferase and GFP are likewise comparable.

Assay Principle

Prior to incubation with protease, the addition of LH_2 and Mg-ATP to the fusion protein substrates produced signals at ~705 nm from SRET and at ~560 nm mainly due to residual bioluminescence (Table 1, Figs. 2f, 4). The change in the ratio of the two peaks produced by specific protease cleavage is the basis of the SRET assays. As a result of severing the linker joining the labeled-RFP and Ppy WT-TS, a decrease in the intensity of the long wavelength peak was expected because the SRET process, which is dependent on energy transfer from the initial Ppy WT-TS emission, is disrupted. Concurrently, the Ppy WT-TS emission should increase as the loss of bioluminescence intensity due to non-radiative resonance energy transfer (Fig. 2c) should no longer occur. These simultaneous spectral changes form the basis of the ratiometric method. The peak ratios and signal separation are determined by efficient intramolecular FRET and modest, but useful, BRET efficiencies. We note that the short wavelength peaks initially

Table 3
Relative protease activities toward fusion protein substrates[a].

	Relative Activity[b] (%)			
Substrate	Factor Xa	Caspase 3	Thrombin	Trypsin
BFS-Xa	11.8	nd	0.02	28.9
BFS-C3	nd	100	nd	1.0
BFS-Th	nd	nd	86.7	24.8
BFS-GS	nd	nd	nd	2.3

[a] Proteases (2 μM) and substrates (2 μM) were incubated at 20 °C and rates of change of peak ratios were calculated as described in detail in Materials and methods.
[b] Values were calculated from the change in peak ratio/min and are expressed as % of the rate obtained with caspase 3 and BFS-C3 (0.577/min), defined as 100%. nd, none detected.

contain a very small contribution from residual RFP fluorescence emphasizing the importance of highly efficient FRET (Fig. 2e).

Proteolytic cleavage of fusion protein substrates

To demonstrate that the expected spectral changes would indeed be produced by proteolysis, we incubated mixtures of substrates and enzymes and monitored the resultant emission spectra produced by the addition of LH_2 and Mg-ATP; the results obtained with factor Xa and BFS-Xa (Fig. 4) are representative. As anticipated, the bioluminescence signal increased as the SRET peak decreased over a 25 min period. Exhaustive protease treatment (Fig. 4 inset) eliminated the SRET signal entirely. Next, we determined that all of the substrates were cleaved by the target proteases at their respective P1 sites (Fig. 1) as evidenced by the identification of masses corresponding to the predicted fragments (Table 2).

Ratiometric assays of protease activity

The rates of protease cleavage of the fusion protein substrates were monitored by recording emission spectra and plotting the change in peak ratios with time. Linear responses were obtained over 15 min (Fig. 5) and for at least 1 h (data not shown) for 58 nM to 465 nM factor Xa with BFS-Xa, 0.41 nM to 8.2 nM caspase 3 with BFS-C3 and 1.0 nM to 20 nM thrombin with BFS-Th. Next we demonstrated that the SRET substrates could be used to monitor protease activity in a convenient endpoint format (Fig. 6). Luminescence peak ratio changes were linear with respect to enzyme concentration over the same concentration ranges used to demonstrate the rates of cleavage (Fig. 5). Moreover, we obtained similar endpoint analysis results (Fig. 6) using a scanning fluorometer even though the initial luminescence 560 nm:705 nm peak ratios were higher because of the lower sensitivity of the PMT detector to the

long wavelength peak. While detection limits of 0.41 nM for caspase 3, 1.0 nM for thrombin, and 58 nM for factor Xa were obtained with a scanning fluorometer, 5- to 10-fold lower protease

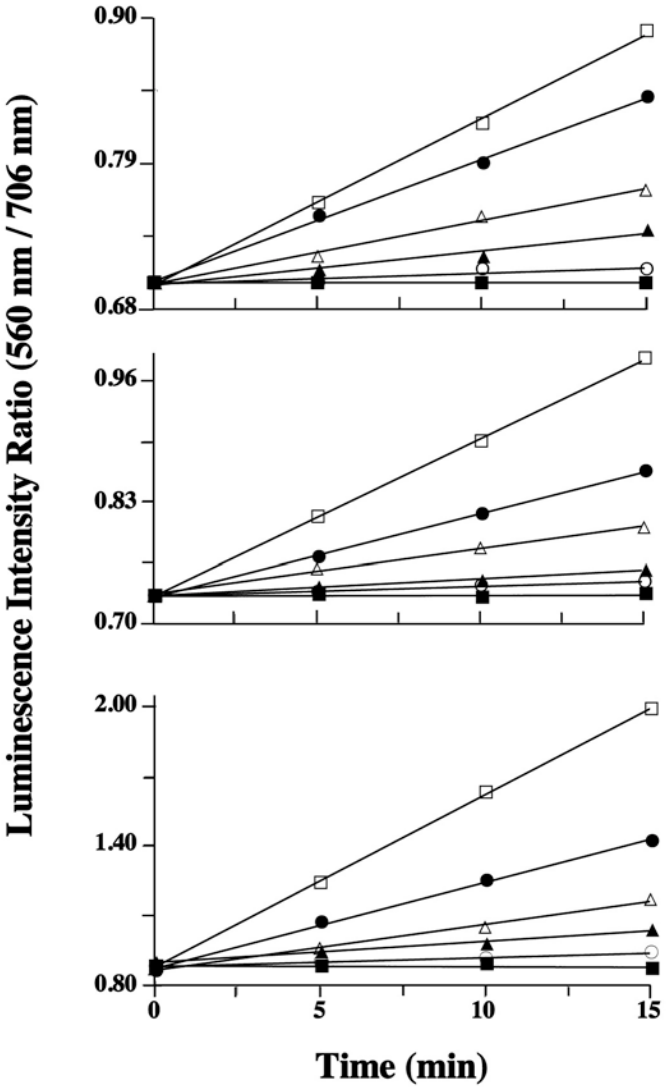


Fig. 5. Rates of protease cleavage of fusion protein substrates monitored by changes in luminescence intensity ratios. Mixtures (0.05 ml) of fusion protein substrates (2.0 μM) and proteases were incubated in PBSA at 20 °C. As described in detail in Materials and methods, aliquots (5 μl) were withdrawn and added to solutions containing LH_2 and Mg-ATP to generate the emission spectra used to calculate 560 nm/ 706 nm peak ratios. Top, BFS-Xa incubated with factor Xa (■, none; ○, 58.0 nM; ▲, 116.0 nM; △, 233.0 nM; ●, 349.0 nM; and □, 465.0 nM); Middle, BFS-C3 incubated with caspase 3 (■, none; ○, 0.408 nM; ▲, 0.817 nM; △, 2.04 nM; ●, 4.08 nM; and □, 8.17 nM); and Bottom, BFS-Th incubated with thrombin (■, none; ○, 1.0 nM; ▲, 2.5 nM; △, 5.0 nM; ●, 10.0 nM; and □, 20.0 nM).

concentrations were detected with an imaging spectrometer equipped with a cooled CCD detector (data not shown). Moreover, the endpoint assays were performed with short incubation

times (15 min), in small volumes (50 µl) using only 2 µM substrate. We note that with FRET-based substrates containing the same protease cleavage sites used in our substrates, detection limits of 0.02 nM and 1 nM have been reported (*19, 20*) for caspase 3 and thrombin, respectively. The assays described here could also be conveniently performed on a simple

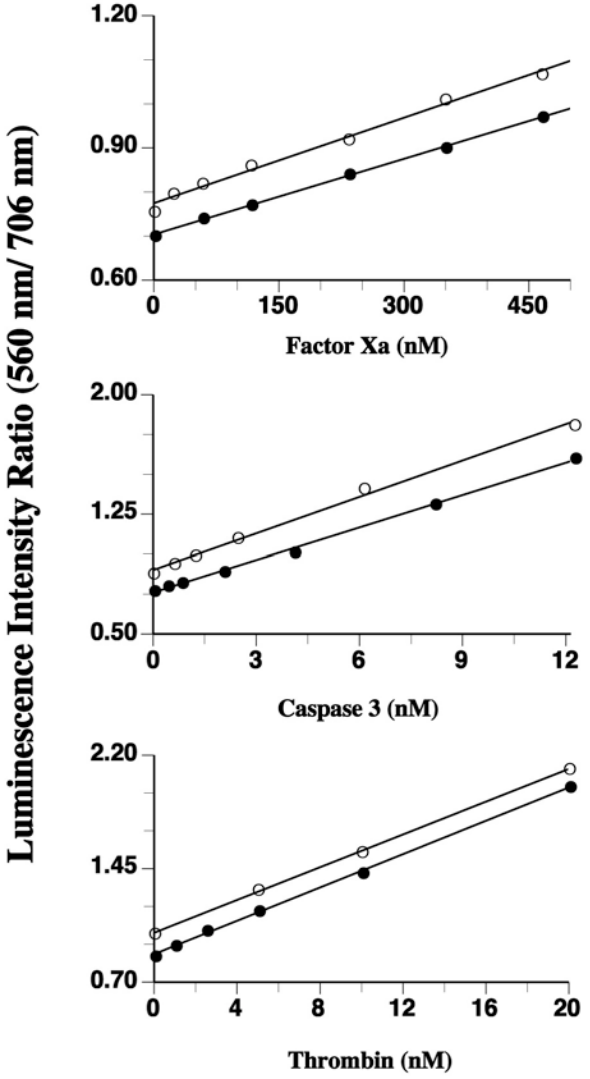


Fig. 6. Relationship between protease concentration and change in luminescence intensity ratio. Mixtures (0.05 ml) of fusion protein substrates and varying concentrations of proteases (BFS-Xa with factor Xa, BFS-C3 with caspase 3 and BFS-Th with thrombin) were incubated in PBSA at 20 °C for 15 min and 5 µl (●) or 10 µl (○) aliquots were withdrawn and added to cuvettes containing 0.50 ml of a 25 mM glycylglycine buffered (pH 7.8) solution of LH_2 (150 µM) and Mg-ATP (2 mM). Samples were briefly mixed and placed in the sample compartment of: (●) a Horiba Jobin-Yvon *i*HR imaging spectrometer or (○) after 1 min into a PerkinElmer LS55 luminescence spectrometer and emission spectra were recorded as detailed in the Materials and methods.

luminometer equipped with two narrow bandpass filters and a red sensitive PMT. Moreover, it also may be possible to monitor protease activity in biological fluids by measuring the disappearance of the SRET peak because of the expected transmission advantage of the nIR signal.

Importantly, the proteases displayed the expected specificity (Table 3) for the consensus tetrapeptide recognition sites contained in the linker portion of the fusion protein substrates (Fig. 1). With the exception of the minor cleavage of BFS-Xa by thrombin, no cross reactivity could be detected and non-specific proteolytic cleavage of the control substrate BFS-GS (Table 3) was not observed.

Conclusions

We have presented results demonstrating for the first time an efficient sequential BRET-FRET energy transfer process based on firefly luciferase bioluminescence. The nIR dye-labeled fusion proteins described here have been employed to assay physiologically important protease activities. The SRET process is the basis for sensitive and specific assays that can be performed in convenient endpoint ratiometric format. With the likely limitation that Cys residues cannot be included in the linker region, it should be possible to readily extend the SRET assays to the measurement of other endoproteases of interest. Future investigations will examine the possibility of altering the linker region to enable assays of additional classes of proteases and to detect analytes through steric disruption of the BRET transfer that is necessary to produce the SRET signals.

References

1. Hastings, J. W. (1996) Chemistries and colors of bioluminescent reactions- a review, *Gene 173*, 5-11.
2. Widder, E. A. (2010) Bioluminescence in the Ocean: Origins of Biological, Chemical, and Ecological Diversity, *Science 328*, 704-708.
3. Roda, A., Guardigli, M., Michelini, E., and Mirasoli, M. (2009) Bioluminescence in analytical chemistry and in vivo imaging, *Trac-Trends in Analytical Chemistry 28*, 307-322.
4. Rowe, L., Dikici, E., and Daunert, S. (2009) Engineering Bioluminescent Proteins: Expanding their Analytical Potential, *Anal. Chem. 81*, 8662-8668.
5. Ando, Y., Niwa, K., Yamada, N., Enomot, T., Irie, T., Kubota, H., Ohmiya, Y., and Akiyama, H. (2008) Firefly bioluminescence quantum yield and colour change by pH-sensitive green emission, *Nature Photonics 2*, 44-47.
6. Branchini, B. R., Ablamsky, D. M., and Rosenberg, J. C. (2010) Chemically Modified Firefly Luciferase Is an Efficient Source of Near-Infrared Light, *Bioconjugate Chem. 21*, 2023-2030.
7. Lakowicz, J. R. (2006) *Principles of Fluorescence Spectroscopy*, Third ed., Springer Science+Business Media, LLC, New York.
8. James, J. R., Oliveira, M. I., Carmo, A. M., Iaboni, A., and Davis, S. J. (2006) A rigorous experimental framework for detecting protein oligomerization using bioluminescence resonance energy transfer, *Nature Methods 3*, 1001-1006.
9. Tsien, R. Y., and Poenie, M. (1986) Fluorescence Ratio Imaging - a New Window into Intracellular Ionic Signaling, *Trends Biochem. Sci 11*, 450-455.
10. Domaille, D. W., Zeng, L., and Chang, C. J. (2010) Visualizing Ascorbate-Triggered Release of Labile Copper within Living Cells using a Ratiometric Fluorescent Sensor, *JACS 132*, 1194-1195.
11. Matveeva, E. G., Gryczynski, Z., Stewart, D. R., and Gryczynski, I. (2010) Ratiometric FRET-based detection of DNA and micro-RNA on the surface using TIRF detection, *J. Lumin. 130*, 698-702.

12. Kiyose, K., Aizawa, S., Sasaki, E., Kojima, H., Hanaoka, K., Terai, T., Urano, Y., and Nagano, T. (2009) Molecular Design Strategies for Near-Infrared Ratiometric Fluorescent Probes Based on the Unique Spectral Properties of Aminocyanines, *Chemistry-a European Journal 15*, 9191-9200.
13. Kikuchi, K. (2010) Design, synthesis and biological application of chemical probes for bio-imaging, *Chem. Soc. Rev. 39*, 2048-2053.
14. Saito, K., Hatsugai, N., Horikawa, K., Kobayashi, K., Matsu-Ura, T., Mikoshiba, K., and Nagai, T. (2010) Auto-luminescent genetically-encoded ratiometric indicator for real-time Ca2+ imaging at the single cell level, *Plos One 5*, e9935.
15. Takeuchi, M., Nagaoka, Y., Yamada, T., Takakura, H., and Ozawa, T. (2010) Ratiometric Bioluminescence Indicators for Monitoring Cyclic Adenosine 3 ',5 '-Monophosphate in Live Cells Based on Luciferase-Fragment Complementation, *Anal. Chem. 82*, 9306-9313.
16. Carriba, P., Navarro, G., Ciruela, F., Ferre, S., Casado, V., Agnati, L., Cortes, A., Mallol, J., Fuxe, K., Canela, E. I., Lluis, C., and Franco, R. (2008) Detection of heteromerization of more than two proteins by sequential BRET-FRET, *Nat. Methods FIELD Full Journal Title:Nature Methods 5*, 727-733.
17. Branchini, B. R., Ablamsky, D. M., Murtiashaw, M. H., Uzasci, L., Fraga, H., and Southworth, T. L. (2007) Thermostable red and green light-producing firefly luciferase mutants for bioluminescent reporter applications, *Anal. Biochem. 361*, 253-262.
18. Neurath, H. (1999) Proteolytic enzymes, past and future, *Proceedings of the National Academy of Sciences of the United States of America 96*, 10962-10963.
19. Zhang, B. (2004) Design of FRET-based GFP probes for detection of protease inhibitors, *Biochem. Biophys. Res. Commun. 323*, 674-678.
20. Boeneman, K., Mei, B. C., Dennis, A. M., Bao, G., Deschamps, J. R., Mattoussi, H., and Medintz, I. L. (2009) Sensing Caspase 3 Activity with Quantum Dot-Fluorescent Protein Assemblies, *JACS 131*, 3828-3829.
21. Gammon, S. T., Villalobos, V. A., Roshal, M., Samrakandi, M., and Piwnica-Worms, D. (2009) Rational Design of Novel Red-Shifted BRET Pairs: Platforms for Real-Time Single-Chain Protease Biosensors, *Biotechnol. Progr. 25*, 559-569.
22. Oliveira, M., Torquato, R. J. S., Alves, M. F. M., Juliano, M. A., Bromme, D., Barros, N. M. T., and Carmona, A. K. (2010) Improvement of cathepsin S detection using a designed FRET peptide based on putative natural substrates, *Peptides 31*, 562-567.
23. Ignjatovic, V., Summerhayes, R., Gan, A., Than, J., Chan, A., Cochrane, A., Bennett, M., Horton, S., Shann, F., Lane, G., Ross-Smith, M., and Monagle, P. (2007) Monitoring Unfractionated Heparin (UFH) therapy: which Anti-Factor Xa assay is appropriate?, *Thromb Res 120*, 347-351.
24. Bialas, A., and Kafarski, P. (2009) Proteases as Anti-Cancer Targets - Molecular and Biological Basis for Development of Inhibitor-Like Drugs Against Cancer, *Anti-Cancer Agents in Medicinal Chemistry 9*, 728-762.
25. Branchini, B. R., Southworth, T. L., Khattak, N. F., Michelini, E., and Roda, A. (2005) Red- and green-emitting firefly luciferase mutants for bioluminescent reporter applications, *Anal. Biochem. 345*, 140-148.
26. Branchini, B. R., Ablamsky, D. M., Rosenman, J. M., Uzasci, L., Southworth, T. L., and Zimmer, M. (2007) Synergistic mutations produce blue-shifted bioluminescence in firefly luciferase, *Biochemistry 46*, 13847-13855.

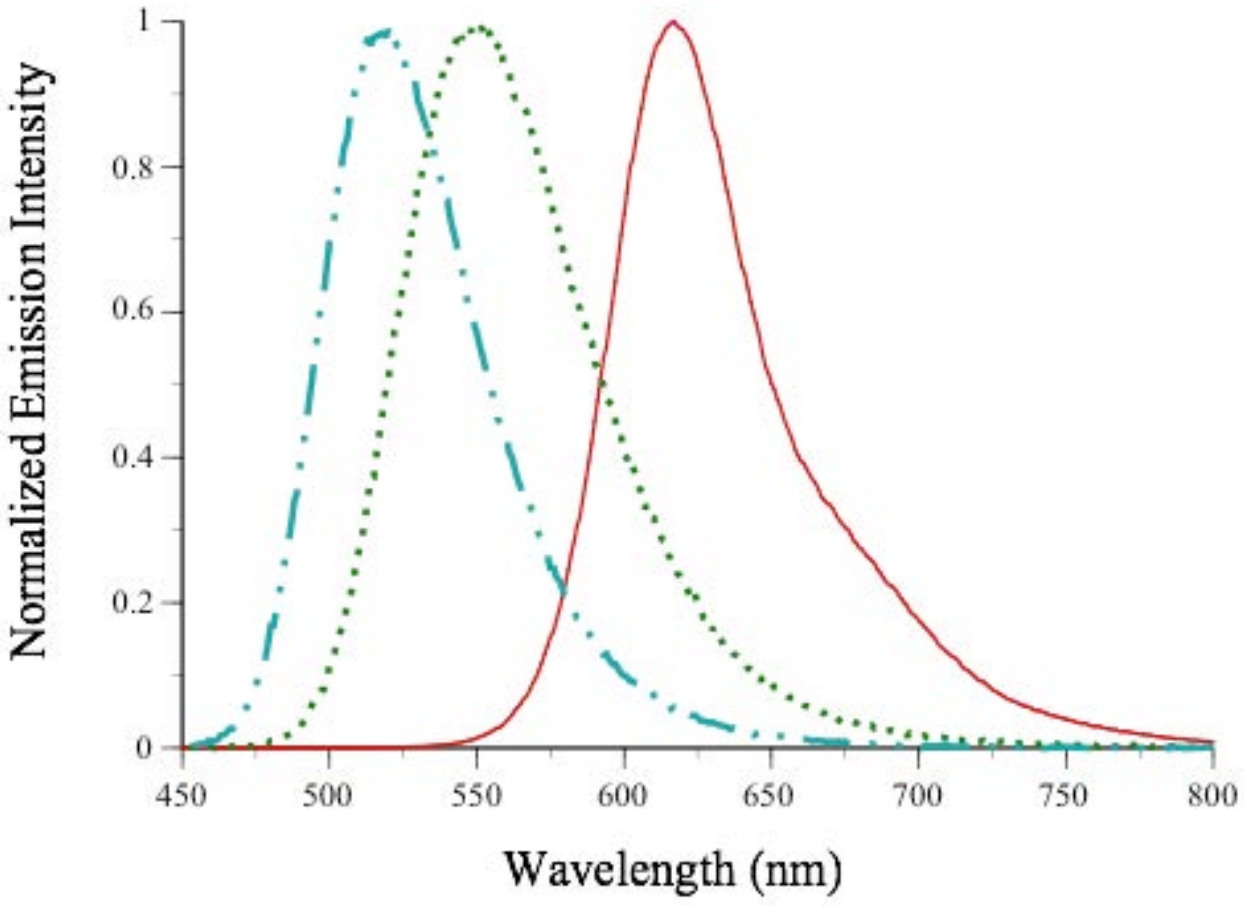
27. Chu, J., Zhang, Z. H., Zheng, Y., Yang, J., Qin, L. S., Lu, J. L., Huang, Z. L., Zeng, S. Q., and Luo, Q. M. (2009) A novel far-red bimolecular fluorescence complementation system that allows for efficient visualization of protein interactions under physiological conditions, *Biosens Bioelectron 25*, 234-239.
28. Haugland, R. P., Spence, M. T. Z., Johnson, I. D., and Basey, A. (2005) *The handbook : a guide to fluorescent probes and labeling technologies*, 10th ed., Molecular Probes, [Eugene, OR].
29. Branchini, B. R., Southworth, T. L., Murtiashaw, M. H., Wilkinson, S. R., Khattak, N. F., Rosenberg, J. C., and Zimmer, M. (2005) Mutagenesis evidence that the partial reactions of firefly bioluminescence are catalyzed by different conformations of the luciferase C-terminal domain, *Biochemistry 44*, 1385-1393.
30. Venkatesh, B., Arifuzzaman, M., Mori, H., Suzuki, S., Taguchi, T., and Ohmiya, Y. (2005) Use of GFP tags to monitor localization of different luciferases in E. coli, *Photochemical & Photobiological Sciences 4*, 740-743.
31. Branchini, B. R., Ablamsky, D. M., Davis, A. L., Southworth, T. L., Butler, B., Fan, F., Jathoul, A. P., and Pule, M. A. (2010) Red-emitting luciferases for bioluminescence reporter and imaging applications, *Anal. Biochem. 396*, 290-297.

Footnotes

[1]*Abbreviations used*: AF680, Alexa Fluor 680 C2-maleimide; BFS-, BRET-FRET substrate, a fusion protein consisting of an N-terminus hexa-His tagged mKate S158A variant labeled with AF680 joined to Ppy WT-TS through the decapeptide linker GSAP4P3P2P1GSG where P4P3P2P1 is DEVD (C3), LVPR (Th), IEGR (Xa) or GSGS (GS); BRET, bioluminescence resonance energy transfer; FRET, fluorescence resonance energy transfer; CCD, charge-coupled device; LC/ESMS, tandem HPLC-electrospray ionization mass spectrometry; LH_2, D-firefly luciferin; Luc, *Photinus pyralis* luciferase (E. C. 1.13.12.7); nIR, near-infrared (700 nm - 1000 nm); PBSA buffer, 20 mM sodium phosphate buffer (pH 7.2) containing 150 mM NaCl, 5 mM EDTA and 0.8 M ammonium sulfate; Ppy WT, recombinant *Photinus pyralis* luciferase containing the additional N-terminal peptide GPLGS; Ppy WT-TS, Ppy WT containing the mutations T214A, A215L, I232A, F295L and E354K; RFP, red fluorescent protein (mKate S158A variant); SRET, sequential BRET-FRET

Major Accomplishment 3:

The results of the initial work with quantum rods and dots are described in publication 7 above. Data are presented below that describe the materials we provided for the initial phase of the project.



Bioluminescence spectra of Ppy RE9 (solid line) and Ppy GR-TS (dotted line) assayed with LH2 and Mg-ATP at pH 7.8 and Ppy GR-TS (dot and dash line) assayed with a LH2 analog and Mg-ATP at pH 9.1.

Properties of luciferases measured at pH 7.8

Enzyme	Relative specific activity		Km (μM)		λ_{max} (nm, fwhm)
	Flash height	Integration (15 min)	LH_2	Mg-ATP	
Ppy WT	100 ± 2	100 ± 2	15 ± 2	160 ± 20	561 (79)
6xHis-Ppy RE9	23 ± 1	50 ± 4	84 ± 5	222 ± 24	617 (57)
6xHis-Ppy GR-TS	39 ± 5	100 ± 4	23 ± 3	163 ± 27	546 (77)
6xHis-Ppy GR-TS with $BtLH_2$	2.3 ± 2	10 ± 10	65 ± 4	-	517 (60)

Materials

The following materials were obtained from the sources indicated: Mg-ATP (bacterial source) from Sigma-Aldrich (St. Louis, MO) and restriction endonucleases from New England Biolabs (Beverly, MA). Firefly luciferin (LH_2)[1] was a generous gift from Promega (Madison, WI). Ppy WT was expressed as a GST-fusion protein and purified by affinity chromatography and stored as described (*1*) in detail previously. The plasmids for 6xHis-Ppy Re9 and 6xHis-Ppy GR-TS were constructed by excising the corresponding genes for Ppy RE9 (*2*) and Ppy GR-TS (*1*) from the pGEX-6P-2 vector and ligating them into a modified pQE30 expression vector using previously described procedures (*3*). The His-tagged proteins were expressed, purified and stored using procedures described (*3*) elsewhere. The found molecular masses (Da) of the proteins not previously reported were within the allowable experimental error (0.01%) of the calculated values (in parenthesis): 6xHis-Ppy RE9, 61 976 (61 974) and 6xHis-Ppy GR-TS, 61 996 (62 002).

General Methods

Protein concentrations were determined with the Bio-Rad Protein Assay system using BSA as the standard. DNA sequencing to verify the ligations was performed at the W. M. Keck Biotechnology Laboratory at Yale University. Specific activity and steady state kinetics measurements were determined as previously reported (*3-5*) except that the final LH_2 concentration was 0.1 mM. Bioluminescence emission spectra were obtained using methods and equipment previously described.(*3*) Mass spectral analyses were performed by tandem HPLC-electrospray ionization mass spectrometry (LC/ESIMS) using a ThermoFinnigan Surveyor HPLC system and a ThermoFinnigan LCQ Advantage mass spectrometer and previously developed conditions for protein mass determinations (*6*).

Footnotes

[1]*Abbreviations used*: 6xHis-, the peptide MetArgGlySerHisHisHisHisHisHisGlySer- replaced the N-terminus GlyProLeuGlySer-; LH_2, D-firefly (beetle) luciferin; 6xHis- ; Ppy GR-TS, Ppy WT containing Thr214Ala/ Ala215Leu/Ile232Ala/Val24Ile/Gly246Ala/Phe250Ser/ Phe295Leu/Glu354Lys; Ppy RE9, Ppy WT containing Thr214Ala/Ala215Leu/Ile232Ala

/Ser284Thr/Phe295Leu/ Arg330Gly/Ile351Val/ Glu354Ile/Phe465Arg ; and Ppy WT, recombinant *Photinus pyralis* luciferase containing the additional N-terminal peptide GlyProLeuGlySer-.

References

1. Branchini, B. R., Ablamsky, D. M., Murtiashaw, M. H., Uzasci, L., Fraga, H., and Southworth, T. L. (2007) Thermostable red and green light-producing firefly luciferase mutants for bioluminescent reporter applications, *Anal. Biochem. 361*, 253-262.

2. Branchini, B. R., Ablamsky, D. M., Davis, A. L., Southworth, T. L., Butler, B., Fan, F., Jathoul, A. P., and Pule, M. A. (2010) Red-emitting luciferases for bioluminescence reporter and imaging applications, *Anal. Biochem. 396*, 290-297.

3. Branchini, B. R., Ablamsky, D. M., and Rosenberg, J. C. (2010) Chemically Modified Firefly Luciferase Is an Efficient Source of Near-Infrared Light, *Bioconjugate Chem. 21*, 2023-2030.

4. Branchini, B. R., Southworth, T. L., Khattak, N. F., Michelini, E., and Roda, A. (2005) Red- and green-emitting firefly luciferase mutants for bioluminescent reporter applications, *Anal. Biochem. 345*, 140-148.

5. Branchini, B. R., Ablamsky, D. M., Rosenman, J. M., Uzasci, L., Southworth, T. L., and Zimmer, M. (2007) Synergistic mutations produce blue-shifted bioluminescence in firefly luciferase, *Biochemistry 46*, 13847-13855.

6. Branchini, B. R., Rosenberg, J. C., Fontaine, D. M., Southworth, T. L., Behney, C. E., and Uzasci, L. (2011) Bioluminescence Is Produced from a Trapped Firefly Luciferase Conformation Predicted by the Domain Alternation Mechanism, *JACS 133*, 11088-11091.

Major Accomplishment 4:

Five novel firefly luciferin analogs were synthesized in which the benzothiazole ring system of the natural substrate was replaced with: benzimidazole, benzofuran, benzothiophene, benzoxazole and indole. The fluorescence, bioluminescence and kinetic properties of the compounds were evaluated with recombinant *Photinus pyralis* wild-type luciferase. With the exception of indole, all of the substrates containing heterocycle substitutions produced readily measurable flashes of light with luciferase. Compared to luciferin, the intensities ranged from 0.3% to 4.4% in reactions with varying pH optima and times to reach maximum intensity. The heteroatom changes influenced both the fluorescence and bioluminescence emission spectra, which displayed maxima of 479 nm to 528 nm and 518 nm to 574 nm, respectively. While there were some interesting trends in the spectroscopic and bioluminescence properties of this group of structurally similar substrate analogs, the most significant findings were associated with the benzothiophene-containing compound. This synthetic substrate produced slow decay glow kinetics that increased the total light based specific activity of luciferase more than 4-fold over the luciferin value. Moreover, over the pH range 6.2 to 9.4 the emission maximum is 523 nm, an unusual 37 nm blue-shift compared to the natural substrate. The extraordinary bioluminescence properties of the benzothiophene luciferin should translate into greater sensitivity for analyte detection in a wide variety of luciferase based applications.

MATERIALS AND METHODS

Materials. The following materials were obtained from the sources indicated: Mg-ATP (bacterial source) from Sigma-Aldrich (St. Louis, MO) and firefly luciferin (LH_2) from Promega (Madison, WI). Recombinant firefly luciferase (PpyWT) was prepared as described previously[18, 19]. Synthetic procedures and characterization of novel compounds are provided in the supporting information.

General Methods. Protein concentrations were determined with the Bio-Rad Protein Assay system using bovine serum albumin as the standard. Ultraviolet (UV)-visible spectra were recorded in 25 mM glycylglycine buffer, pH 7.8, with a Perkin Elmer Lambda 25 spectrometer. Enzyme activity assays were performed with PpyWT, and all data were replicated in triplicate and are reported as the mean ± standard deviation.

Luciferase Activity Assays. Bioluminescence activity assays were performed with a custom-built luminometer assembly containing a Hamamatsu R928 PMT and a C6271 HV power supply socket assembly that was described in detail previously_ENREF_20[20]. Reactions (0.525 mL final volume) were initiated by the injection of 0.120 mL of 9.0 mM ATP into 8 x 50-mm polypropylene tubes containing 0.4 mL of buffer with luciferin or substrate analogs and 5 mL (0.4 – 4 μg) of enzyme in 20 mM Tris-HCl (pH 7.4 at 4 °C) containing 150 mM NaCl, 1 mM EDTA, 1 mM DTT, 0.8 M ammonium sulfate and 2% glycerol. Assay buffers were: 25 mM glycylglycine, pH 8.2; and 50 mM 2-amino-2-methyl-1,3-propanediol (AMPD), pH 8.5 or pH 9.1 for $BtLH_2$. The final LH_2 and analog concentrations were ~6 times their K_m values. Peak height and integrated intensity values were recorded and corrected for the spectral response of the detector.

Bioluminescence Emission Spectra. Emission spectra were obtained using a Horiba Jobin-Yvon *i*HR imaging spectrometer equipped with a liquid N_2 cooled CCD detector and the excitation source turned off. Data were collected at 25 °C in a 0.8 ml quartz cuvette over the

wavelength range 400–935 nm with the emission slit width set to 10 nm. Reactions (0.525 mL final volume) were initiated by addition of 5 mL of PpyWT stock solution (15 - 150 nM final concentration) to cuvettes containing substrate (150 mM) and Mg-ATP (2 mM) in 25 mM 2[N-morpholino]ethanesulfonic acid (MES), pH 6.2; 25 mM 4-(2-hydroxyethyl)-1-piperazineethane-sulfonic acid (HEPES), pH 7.0; 25 mM glycylglycine buffer, pH 7.8 or 8.2; 50 mM AMPD, pH 8.5, 8.6, 9.1 or 9.4. The pH values were confirmed before and after spectra were obtained. All spectra were corrected for the spectral response of the CCD using a correction curve provided by the manufacturer.

Relative Bioluminescence Quantum Yields. Bioluminescent quantum yields for LH_2 and the substrate analogs were determined from bioluminescence activity assays in which a limiting amount of compound was reacted with an excess of luciferase under single turnover conditions. Into 0.4 mL of buffer at the optimal pH for each compound (see above) containing 65 nM substrate and 3.25 *m*M PpyWT, 0.12 mL of 9 mM Mg-ATP in the same buffer was injected. The light output was monitored until the initial signal intensity decreased by 99% at a sampling rate of 100 Hz. An additional aliquot of enzyme was added to the spent mixtures and emission intensity was monitored to ensure that the reactions were completed. The total integrated light intensities were reported relative to the value obtained with LH_2 at pH 8.2.

Steady-State Kinetic Constants. Values of K_m and V_{max} for the substrate analogs were determined as previously described[21] from bioluminescence activity assays in which measurements of maximal light intensities were taken as estimates of initial velocities. Activity assays were performed at the pH optimum for each compound (Table 1). Briefly, the measurements were made with Mg-ATP at a saturating level (2 mM) and varying concentrations of substrate: LH_2, 2 – 500 μM; $BtLH_2$, 3 -750 μM; $BoLH_2$, 0.6 – 1.0 mM; $BiLH_2$, 5 – 200 μM and $BfLH_2$, 5 – 600 μM. Data were collected and analyzed using Enzyme Kinetics Pro software (Syntex, Palo Alto, CA).

RESULTS AND DISCUSSION

Luciferin analog synthesis and characterization. The novel luciferin analogs were synthesized as shown in Scheme 2. All were accessed via reaction of D-cysteine with the corresponding nitrile precursor under mild aqueous conditions. While the appropriate precursor for benzoxazole luciferin $BoLH_2$ was prepared directly by treating aminoresorcinol with Appel's salt[22], the other analogs made use of a phenolic methyl ether protecting group which was removed by brief treatment with molten pyridinium hydrochloride. Synthesis of the indole core has been previously reported and was achieved via a similar route[23]. The nitrile intermediates for the benzofuran and benzothiophene luciferins were approached from a common 4-methoxysalicylaldehyde precursor. In the case of the benzofuran luciferin, direct reaction with a haloacetonitrile and subsequent condensation of the activated methylene with the proximal aldehyde afforded the methoxy benzofuranonitrile. The latter transformation was somewhat sensitive to reaction conditions, as the nitrile was prone to hydration to form the amide; however potassium carbonate in anhydrous DMF at high temperature proved quite effective[24]. The benzothiophene core was accessed by thiocarbamoylation of the same salicylaldehyde, followed by a heat-induced rearrangement[25] and subsequent hydrolysis to furnish the thiophenol. The conditions used to alkylate the thiophenol intermediate resulted in spontaneous condensation of the aldehyde with the activated methylene to yield the desired benzothiophene nitrile. Finally,

the benzimidazole derivative $BiLH_2$ was synthesized from the commercially available 2-thiobenzimidazole via methylation, oxidation to the sulfone, and subsequent displacement with potassium cyanide. The methoxybenzimidazole-2-carbonitrile was elaborated to $BiLH_2$ by standard deprotection and cyclization with D-cysteine.

As expected[16], the H1- and C13-NMR spectra of the $BiLH_2$ analog displayed peak broadening of the resonances due to rapid tautomerization of the benzimidazole ring in solution. It was not possible to determine to what extent $BiLH_2$ exists as the 5´ or 6´- hydroxy substituted compound at the luciferase active site. While this ambiguity is problematic, especially because the 5´-substituted regioisomer is expected to be inactive in light production, the results and interpretations that follow were made using the assumption that $BiLH_2$ exists solely as the 6´-hydroxy substituted analog.

Kinetic Properties with PpyWT. Prior to determining the performance of the substrate analogs with PpyWT, we first determined the optimum assay pH (Table 1) based on the intensity of the flash height produced with Mg-ATP. We then compared the properties of the compounds at their respective pH optima, which ranged from 8.2 to 9.1, throughout the rest of the study. Little variation in the K_m values was observed with the exception that the value for $BoLH_2$ was unexpectedly 6-fold higher than that of LH_2. The parameters K_m, k_{cat} and k_{cat}/K_m were determined from assays performed at the respective pH optima (Table 1) in which the peak intensities were taken as measures of initial velocity. For PpyWT the highest k_{cat} and k_{cat}/K_m values, measures of the catalytic efficiency of PpyWT with the analogs, were obtained with $BtLH_2$ (Table 2). Additionally, the emission kinetics (rise and decay times) varied significantly (Table 2). While LH_2, $BoLH_2$ and $BiLH_2$ with Y = N (Figure 1), reached maximum intensity from 0.35 to 0.6 s, $BtLH_2$ and $BfLH_2$ with Y = CH had rise times of 6 s and 4 s, respectively. $BtLH_2$ had the most sustained emission as indicated by a decay time of 95 min greatly exceeding the 23 s value of LH_2. This "glow" type kinetic profile is unusual for PpyWT, which typically displays strong product inhibition and flash kinetics, and contributed to the high relative bioluminescence quantum yield of $BtLH_2$ as discussed below.

Spectral Emission Properties. The fluorescence emission spectra (Figure 2) of the substrate analogs were recorded using a common excitation wavelength of 330 nm; the fluorescence emission maxima and quantum yields (Φ_{Fl}) are summarized in Table 1. The emission maxima varied from 429 nm ($InLH_2$) to 537 nm (LH_2) and the quantum yields (measured at pH 11.0) from 0.01 ($InLH_2$) to 0.83 (LH_2). With respect to emission maxima, the substrates fell into two groups: (1) LH_2, $BoLH_2$, and $BiLH_2$ that have λ_{max} = ~530 nm (range 528 nm-537 nm) and (2) $InLH_2$, $BfLH_2$ and $BtLH_2$ that have λ_{max} values in the range 429 nm to 485 nm. The clear distinguishing factor is the absence or presence of a nitrogen atom (Y = N in Figure 1) that is associated with longer wavelength emission.

The range of bioluminescence emission maxima (Table 3 and Figure 3) at the respective pH optima for the substrate-luciferase reactions varied 56 nm from 518 nm ($BfLH_2$) to 574 nm ($BiLH_2$). When Y = CH, green emission (518 nm – 523 nm) is observed for $BtLH_2$, and $BfLH_2$. Yellow-green light (556 nm – 574 nm) was observed with LH_2, $BiLH_2$, and $BoLH_2$ that have Y = N. Data could not be obtained for $InLH_2$ because of extremely low bioluminescent activity, and this analog was not further evaluated.

The spectral shift of approximately 50 nm over the pH range 6.2 to 9.4 with LH_2 is an example of the pH sensitivity of the true firefly luciferases (Table 3). Only $BoLH_2$ and $BiLH_2$ show very similar pH behavior to LH_2, while $BtLH_2$ and $BfLH_2$ are almost insensitive to pH change. Apparently, fluorescence and bioluminescence emission as well as pH sensitivity and

reaction rise times are related to whether Y = CH or Y = N. In the crystal structures of *L. cruciata* and *P. pyralis* luciferases[26, 27] in complex with an inhibitor structurally similar to L-AMP, the N atom at the Y position of the inhibitor is hydrogen bonded to S347 through an intervening water molecule at the active site. A disruption of this interaction may be responsible for the resistance to red-shifting emission at low pH, the long wavelength bioluminescence and slow rise times.

Relative Bioluminescence Quantum Yields (Φ_{Bl}). The Φ_{Bl} values (relative to the value obtained with LH_2) were measured with excess enzyme under single turnover conditions (Table 1). The results also fall into two groups: (1) those with high (≥70%) values (LH_2, $BtLH_2$) and (2) those with very low (≤14%) values ($BiLH_2$, $BfLH_2$, $BoLH_2$, and $InLH_2$). Here there is a correlation with the electronegativity of the X position atom (Figure 1). Higher values are associated with the low electronegativity of S, while O and N produce lower Φ_{Bl} values. This effect is especially powerful as one might expect that the Φ_{Bl} values would follow the fluorescence quantum yield (Φ_{Fl}) trend (Table 1) suggesting, for example, that $BoLH_2$ (Φ_{Fl} = 0.71) should have the highest Φ_{Bl} value among the analogs. Instead, $BoLH_2$ produced little light (~1% compared to LH_2). Evidently, the high electronegativity of the oxygen atom interferes with the efficient formation of the excited state analog oxyluciferins (Φ_{ES}) and/or causes the corresponding adenylates to form dehydroluciferin-like product in great excess over oxyluciferin (Φ_{Rx}). Possibly too, adenylate formation is severely reduced (Scheme 1). In marked contrast, $BtLH_2$ efficiently produces light with a relative Φ_{Bl} equal to 70% that of LH_2 despite having a Φ_{Fl} = 0.33, only 40% of the LH_2 value. This analog may be capable of producing a greater percentage of analog oxyluciferin in the excited state (Φ_{ES}) and/or with a higher analog oxyluciferin/L-AMP ratio (Φ_{Rx}). The latter is likely occurring as evidenced by the extended decay when substrate is in excess. The slow decay may be due to slower PpyWT inhibition because less analog L-AMP is formed and/or because it has a higher K_i value than L-AMP. Lower product inhibition by the oxyluciferin analog may also contribute. These arguments, like those above, assume that the heterocyclic analogs bind in the same relative orientation as LH_2 with respect to rotation about the C2-C2´ bond. Otherwise, the relative position of the 6´-hydroxyl at the active site will be altered and would likely diminish light emission.

$BtLH_2$ is a promising Luc substrate. Firefly luciferase assays based on ATP detection, reporter gene detection and in vivo bioluminescence imaging can be performed more conveniently and with greater sensitivity when the signal decays slowly displaying "glow" rather than "flash" kinetics. Interestingly, although the initial intensity of $BtLH_2$ is only ~4% of that produced with LH_2, this intensity decays so slowly (Figure 4) that over 1.5 hours, ~4-fold greater total light is emitted than with the natural substrate (Table 2). To demonstrate the importance of the sustained light emission of $BtLH_2$, we recorded images of reactions with identical amounts of PpyWT and Mg-ATP and saturating concentrations of substrates under optimal conditions, pH 8.2 for LH_2 and pH 9.1 for the analog. Initially the LH_2-containing reaction is brighter, but after 5 min the reactions are of approximately equal intensity (Figure 4 inset). However, the bright emission with $BtLH_2$, but not LH_2, persisted even after 30 min (Figure 4 inset). Also, the blue shifted emission of PpyWT with $BtLH_2$ will provide better signal separation with red bioluminescence signals than is possible with LH_2. Additionally, we evaluated the potential of $BtLH_2$ for imaging studies by using it to visualize single *E. coli* colonies expressing PpyWT (Figure 5). Surprisingly, despite the higher pH optimum of bioluminescence with $BtLH_2$, the signal intensities produced in bacteria with the analog and LH_2 are very similar.

We have developed synthetic methods for the preparation of five novel LH_2 analogs. Among them, $BtLH_2$ displays the most promise as an alternative substrate for Luc in application where it is advantageous to have blue-shifted and longer lived emission than is typically observed with the natural substrate. These attributes have been achieved without the necessity of chemical additives. Studies are in progress to develop firefly luciferases that selectively enhance the properties of $BtLH_2$ compared to the natural substrate.

REFERENCES

1. Ribeiro, C., and da Silva, J. (2008) Kinetics of inhibition of firefly luciferase by oxyluciferin and dehydroluciferyl-adenylate, *Photochemical & Photobiological Sciences 7*, 1085-1090.
2. Ando, Y., Niwa, K., Yamada, N., Enomot, T., Irie, T., Kubota, H., Ohmiya, Y., and Akiyama, H. (2008) Firefly bioluminescence quantum yield and colour change by pH-sensitive green emission, *Nat Photonics 2*, 44-47.
3. White, E. H., Rapaport, E., Seliger, H. H., and Hopkins, T. A. (1971) Chemi- and bioluminescence of firefly luciferin. Efficient chemical production of electronically excited states, *Bioorg. Chem. 1*, 92-122.
4. White, E. H., Worther, H., Field, G. F., and McElroy, W. D. (1965) Analogs of Firefly Luciferin, *J. Org. Chem. 30*, 2344 - 2348.
5. White, E. H., and Worther, H. (1966) Analogs of firefly luciferin. 3, *J Org Chem 31*, 1484-1488.
6. White, E. H., Worther, H., Seliger, H. H., and McElroy, W. D. (1966) Amino analogs of firefly luciferin and biological activity thereof, *JACS 88*, 2015-2019.
7. Branchini, B. R., Hayward, M. M., Bamford, S., Brennan, P. M., and Lajiness, E. J. (1989) Naphthylluciferin and Quinolylluciferin - Green and Red-Light Emitting Firefly Luciferin Analogs, *Photochem. Photobiol. 49*, 689-695.
8. Farace, C., Blanchot, B., Champiat, D., Couble, P., Declercq, G., and Millet, J. L. (1990) Synthesis and characterization of a new substrate of Photinus pyralis luciferase: 4-methyl-D-luciferin, *J Clin Chem Clin Biochem 28*, 471-474.
9. Branchini, B. R. (2000) Chemical synthesis of firefly luciferin analogs and inhibitors, *Method Enzymol 305*, 188-195.
10. Shinde, R., Perkins, J., and Contag, C. H. (2006) Luciferin derivatives for enhanced in vitro and in vivo bioluminescence assays, *Biochemistry 45*, 11103-11112.
11. Woodroofe, C. C., Shultz, J. W., Wood, M. G., Osterman, J., Cali, J. J., Daily, W. J., Meisenheimer, P. L., and Klaubert, D. H. (2008) N-alkylated 6 '-aminoluciferins are bioluminescent substrates for Ultra-Glo and QuantiLum luciferase: New potential scaffolds for bioluminescent assays, *Biochemistry 47*, 10383-10393.
12. Reddy, G. R., Thompson, W. C., and Miller, S. C. (2010) Robust Light Emission from Cyclic Alkylaminoluciferin Substrates for Firefly Luciferase, *JACS 132*, 13586-13587.
13. Takakura, H., Sasakura, K., Ueno, T., Urano, Y., Terai, T., Hanaoka, K., Tsuboi, T., and Nagano, T. (2010) Development of Luciferin Analogues Bearing an Amino Group and Their Application as BRET Donors, *Chemistry-an Asian Journal 5*, 2053-2061.

14. Takakura, H., Kojima, R., Urano, Y., Terai, T., Hanaoka, K., and Nagano, T. (2011) Aminoluciferins as Functional Bioluminogenic Substrates of Firefly Luciferase, *Chemistry-an Asian Journal 6*, 1800-1810.
15. Conley, N. R., Dragulescu-Andrasi, A., Rao, J. H., and Moerner, W. E. (2012) A Selenium Analogue of Firefly D-Luciferin with Red-Shifted Bioluminescence Emission, *Angewandte Chemie-International Edition 51*, 3350-3353.
16. McCutcheon, D. C., Paley, M. A., Steinhardt, R. C., and Prescher, J. A. (2012) Expedient Synthesis of Electronically Modified Luciferins for Bioluminescence Imaging, *JACS 134*, 7604-7607.
17. White, E. H., McCapra, F., and Field, G. F. (1963) The structure and synthesis of firefly luciferin, *JACS 85*, 337-343.
18. Branchini, B. R., Ablamsky, D. M., Murtiashaw, M. H., Uzasci, L., Fraga, H., and Southworth, T. L. (2007) Thermostable red and green light-producing firefly luciferase mutants for bioluminescent reporter applications, *Anal. Biochem. 361*, 253-262.
19. Branchini, B. R., Ablamsky, D. M., Davis, A. L., Southworth, T. L., Butler, B., Fan, F., Jathoul, A. P., and Pule, M. A. (2010) Red-emitting luciferases for bioluminescence reporter and imaging applications, *Anal. Biochem. 396*, 290-297.
20. Branchini, B. R., Ablamsky, D. M., Rosenman, J. M., Uzasci, L., Southworth, T. L., and Zimmer, M. (2007) Synergistic mutations produce blue-shifted bioluminescence in firefly luciferase, *Biochemistry 46*, 13847-13855.
21. Branchini, B. R., Southworth, T. L., Murtiashaw, M. H., Wilkinson, S. R., Khattak, N. F., Rosenberg, J. C., and Zimmer, M. (2005) Mutagenesis evidence that the partial reactions of firefly bioluminescence are catalyzed by different conformations of the luciferase C-terminal domain, *Biochemistry 44*, 1385-1393.
22. Cuadro, A. M., and Alvarezbuilla, J. (1994) 4,5-Dichloro-1,2,3-Dithiazolium Chloride (Appels Salt) - Reactions with N-Nucleophiles, *Tetrahedron 50*, 10037-10046.
23. Borza, I., Bozo, E., Barta-Szalai, G., Kiss, C., Tarkanyi, G., Demeter, A., Gati, T., Hada, V., Kolok, S., Gere, A., Fodor, L., Nagy, J., Galgoczy, K., Magdo, I., Agai, B., Fetter, J., Bertha, F., Keseru, G. M., Horvath, C., Farkas, S., Greiner, I., and Domany, G. (2007) Selective NR1/2B N-Methyl-d-aspartate receptor antagonists among indole-2-carboxamides and benzimidazole-2-carboxamides, *J. Med. Chem. 50*, 901-914.
24. Hirota, T., Fujita, H., Sasaki, K., Namba, T., and Hayakawa, S. (1986) A Novel Synthesis of Benzofuran and Related-Compounds .1. The Vilsmeier Reaction of Phenoxyacetonitriles, *J. Heterocycl. Chem. 23*, 1347-1351.
25. Kolmakov, K. A., and Kresge, A. J. (2008) Synthesis of possible o-thioquinone methide precursors, *Can J Chem 86*, 119-123.
26. Nakatsu, T., Ichiyama, S., Hiratake, J., Saldanha, A., Kobashi, N., Sakata, K., and Kato, H. (2006) Structural basis for the spectral difference in luciferase bioluminescence, *Nature 440*, 372-376.
27. Sundlov, J. A., Fontaine, D. M., Southworth, T. L., Branchini, B. R., and Gulick, A. M. (2012) Crystal Structure of Firefly Luciferase in a Second Catalytic Conformation Supports a Domain Alternation Mechanism, *Biochemistry 51*, 6493-6495.

Table 1. Fluorescence and bioluminescence properties of LH_2 and analogs

substrate	pK_a[a]	pH optima[b]	fluorescence λ_{max}[c]	fluorescence Φ_{Fl}[d] pH 11.0	fluorescence Φ_{Fl}[d] pH optima	relative bioluminescence Φ_{Bl}
LH_2	8.6	8.2	537	0.83	0.71	100 ± 1.5
$BtLH_2$	9.1	9.1	485	0.33	0.28	70 ± 2.5
$BoLH_2$	8.8	8.5	528	0.71	0.57	0.9 ± 0.1
$BiLH_2$	9.6	8.5	533	0.03	0.02	14 ± 1.0
$BfLH_2$	9.2	8.5	479	0.08	0.07	1.9 ± 0.1
$InLH_2$	10.0	ND[e]	429	0.02	0.01	< 0.1%

[a]pK_a values are within error of ± 0.1. [b]pH optima of reactions with PpyWT values are within error of ± 0.1. [c]Fluorescence λ_{max} values are within error of ± 1. [d]Fluorescence quantum yields were measured at pH 11.0 and reaction pH optima with 370 nm excitation and values are within error of ± 5%. [e]Data could not be obtained for $InLH_2$ because of low bioluminescent activity.

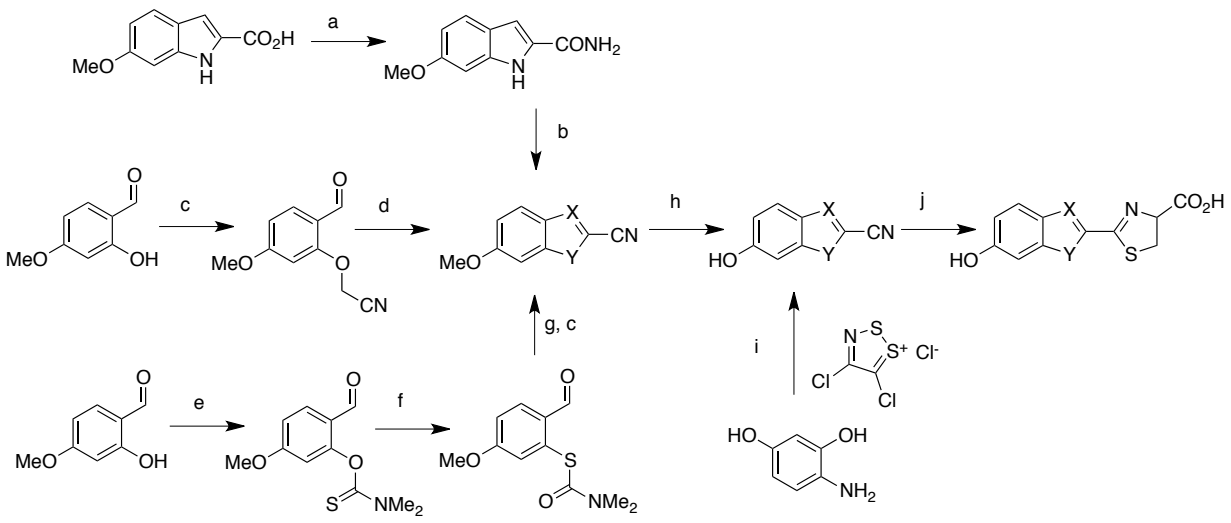
Table 2. Kinetic properties of LH_2 and analogs[a]

substrate	K_m (μM)	k_{cat}[b] (s^{-1})	k_{cat}/K_m ($mM^{-1}\ s^{-1}$)	rise time[c] (s)	decay time to 10%[c] (s)	relative specific activity[d] flash height	relative specific activity[d] integration
LH_2	28 ± 2	1.9E-01	6.78 ± 0.14	0.36	23	100	100
$BtLH_2$	61 ± 2	6.7E-03	0.11 ± 0.002	6	5700	4.4	421
$BoLH_2$	176 ± 26	1.8E-03	0.01 ± 0.001	0.35	500	1.0	13
$BiLH_2$	20 ± 5	5.1E-04	0.03 ± 0.001	0.6	400	0.3	0.4
$BfLH_2$	70 ± 4	6.3E-04	0.01 ± 0.001	4	150	0.6	1.7

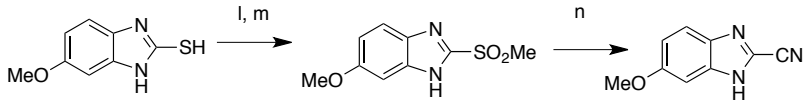
[a]Data could not be obtained for $InLH_2$ because of low bioluminescent activity. [b]k_{cat} values are within error of ± 5%. [c]Rise (time to reach maximum intensity) and decay times are within error of ± 10%. [d]Relative specific activity values are within error of ± 10%.

Table 3. Bioluminescence emission spectra[a]

substrate	bioluminescence λ_{max} (nm) pH 6.2	pH 7.0	pH 7.8	pH 8.6	pH 9.4	pH optima
LH_2	613 (65)	562 (89)	561 (73)	560 (68)	559 (69)	560 (71)
$BtLH_2$	525 (61)	522 (56)	523 (57)	523 (57)	523 (57)	523 (57)
$BoLH_2$	604 (63)	580 (94)	561 (86)	557 (75)	557 (71)	557 (77)
$BiLH_2$	614 (96)	600(103)	577 (96)	574 (80)	570 (78)	574 (80)
$BfLH_2$	ND[b]	527 (47)	519 (57)	518 (58)	516 (62)	518 (57)

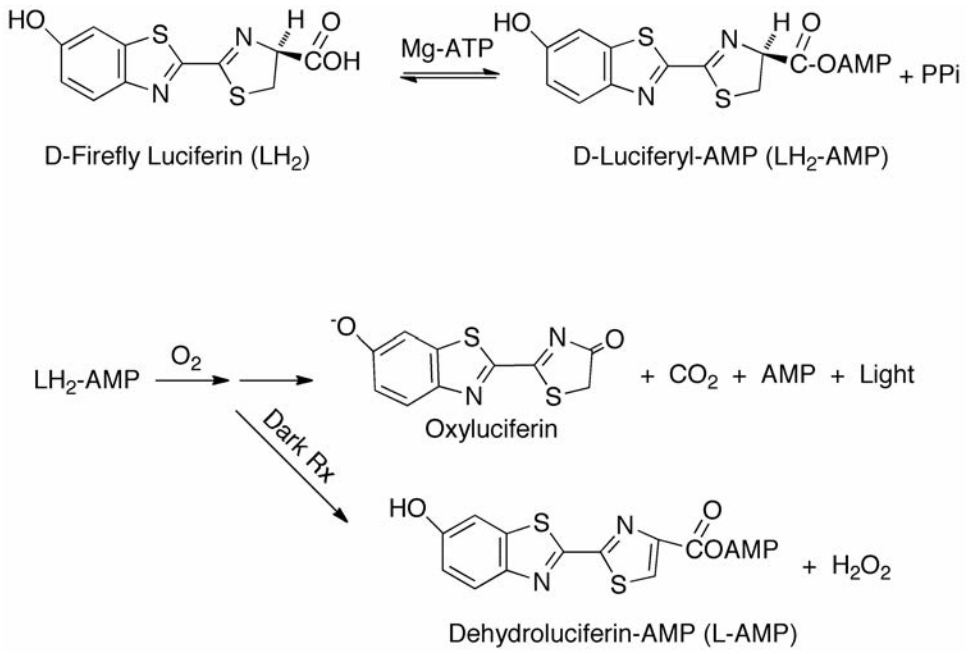


a. i. $SOCl_2$, DMF(cat) ii. NH_4OH, H_2O b. $POCl_3$ c. $BrCH_2CN$, K_2CO_3, THF d. K_2CO_3, DMF, 100 C e. $ClC(S)NMe_2$, DBU, DMF 0 C. f. PhMe, 180 C. g. NaOH, H_2O h. pyridine HCl, 220 C. i. THF/DCM, 0 C, then 150 C 10 min. j. D-cysteine, H_2O/MeCN.



l. MeI, TEA. m. mCPBA. n. KCN, DMSO.

Scheme 2. Synthetic Pathways to Luciferin Analogs



Scheme 1

FIGURE LEGENDS

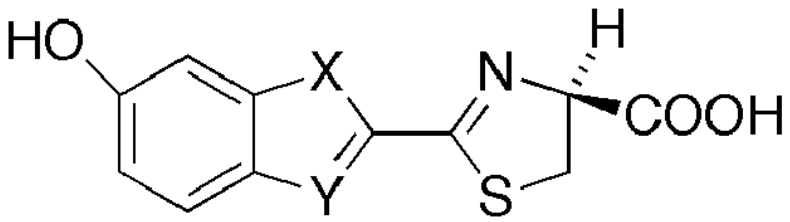
Figure 1. Chemical structures of firefly luciferin (LH_2) and substrate analogs.

Figure 2. Normalized fluorescence emission spectra of LH_2 (black) and analogs ($InLH_2$, cyan; $BfLH_2$, green; $BtLH_2$, blue; $BoLH_2$, orange; $BiLH_2$, magenta) recorded at the pH optima for bioluminescence (Table 1). Spectra were acquired with 330 nm excitation as described in detail in the Supporting Information.

Figure 3. Normalized bioluminescence emission spectra of LH_2 (black) and the analogs ($BfLH_2$, green; $BtLH_2$, blue; $BoLH_2$, orange; $BiLH_2$, magenta) recorded at the pH optima for bioluminescence (Table 1). Additional experimental details are described in the Materials and Methods.

Figure 4. Light emission time courses for PpyWT catalyzed bioluminescence reactions of LH_2 at pH 8.2 and $BtLH_2$ at pH 9.1 recorded as described in Materials and Methods. Top right inset: Bioluminescence emission images of in vitro reactions of LH_2 (top) and $BtLH_2$ (bottom) taken (A) 5 s, (B) 5 min, and (C) 30 min after initiation of light reactions. Assays (0.130 mL volume) in 25 mM glycylglycine buffer, pH 8.2 for LH_2, or 50 mM AMPD, pH 9.1 for $BtLH_2$, contained 2.0 mM Mg-ATP, 1 μg of enzyme, and substrate concentrations ~6 times their K_m values. All images were obtained with a ChromaScan Lite Imaging System.

Figure 5. Bioluminescence imaging of *E. coli* colonies on nitrocellulose filters expressing PpyWT. Light reactions were initiated by soaking nitrocellulose filters with 0.6 mL of 1 mM solutions of LH_2 (left) and $BtLH_2$ (right) in 0.1 M sodium citrate buffer, pH 5.5. After ~3 min, images were obtained with a ChromaScan Lite Imaging System.



LH_2: X = S, Y = N
$BiLH_2$: X = NH, Y = N
$BoLH_2$: X = O, Y = N
$BtLH_2$: X = S, Y = CH
$InLH_2$: X = NH, Y = CH
$BfLH_2$: X = O, Y = CH

Figure 1

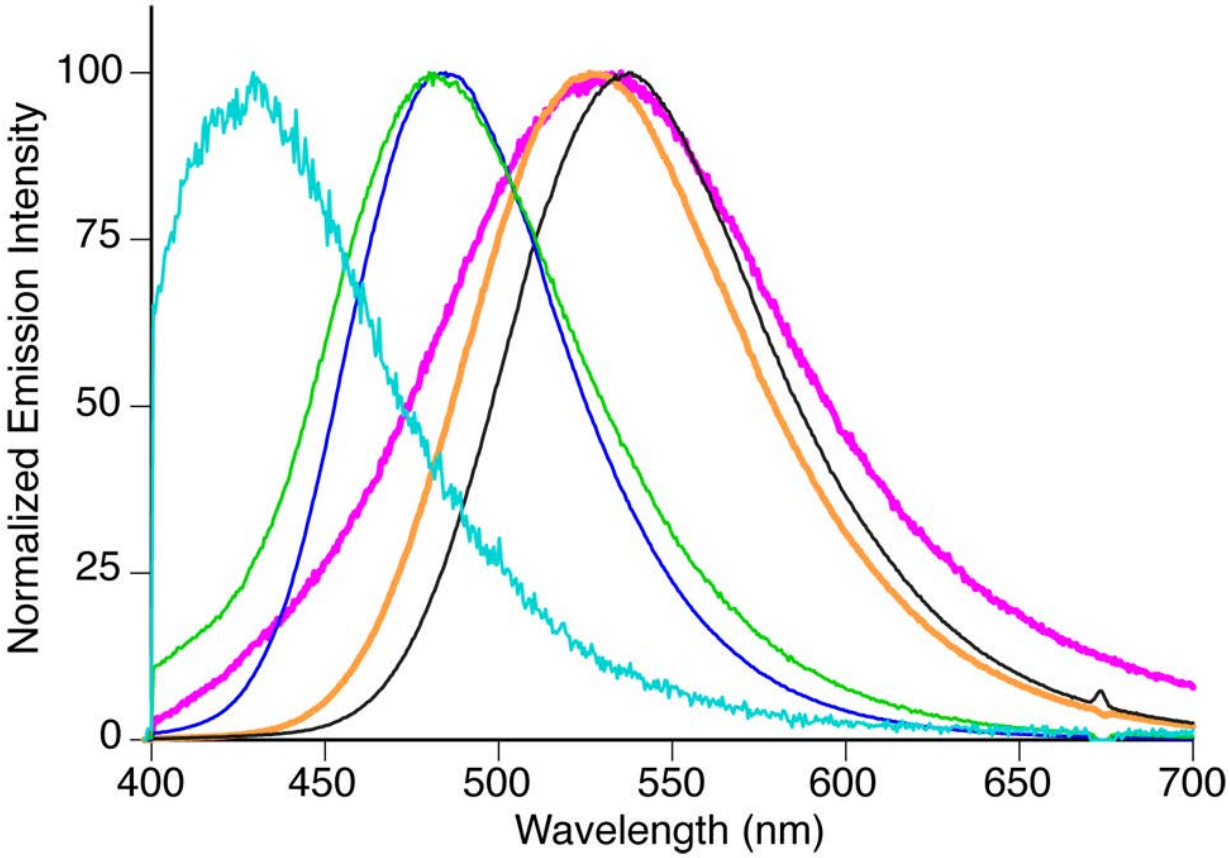


Figure 2

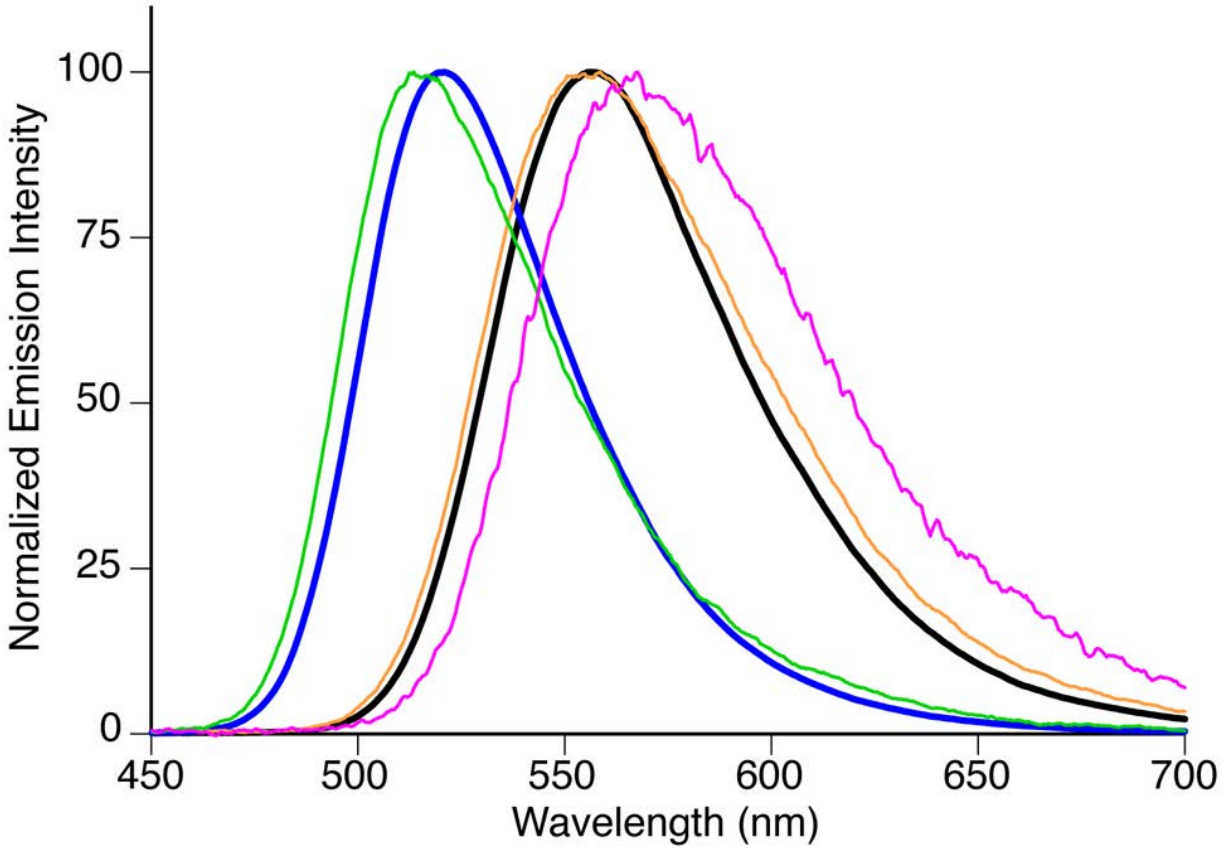

Normalized Emission Intensity
100
75
50
25
0
450
500
550
600
650
700
Wavelength (nm)

Figure 3

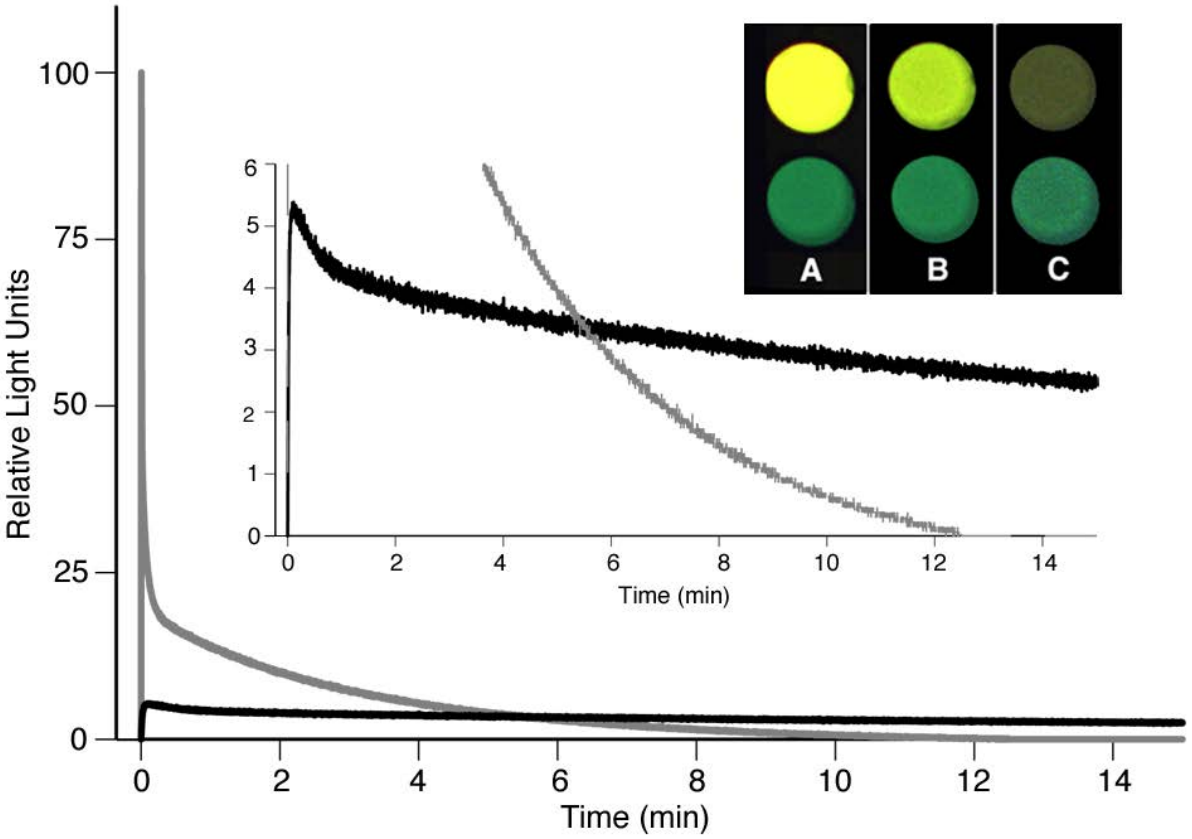


Figure 4

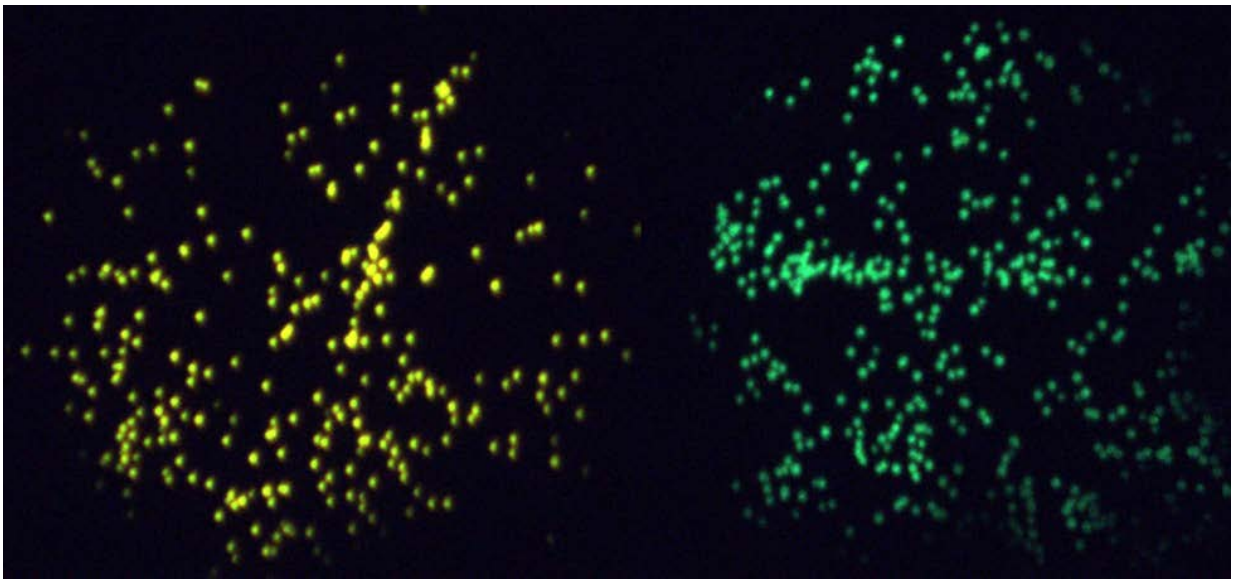
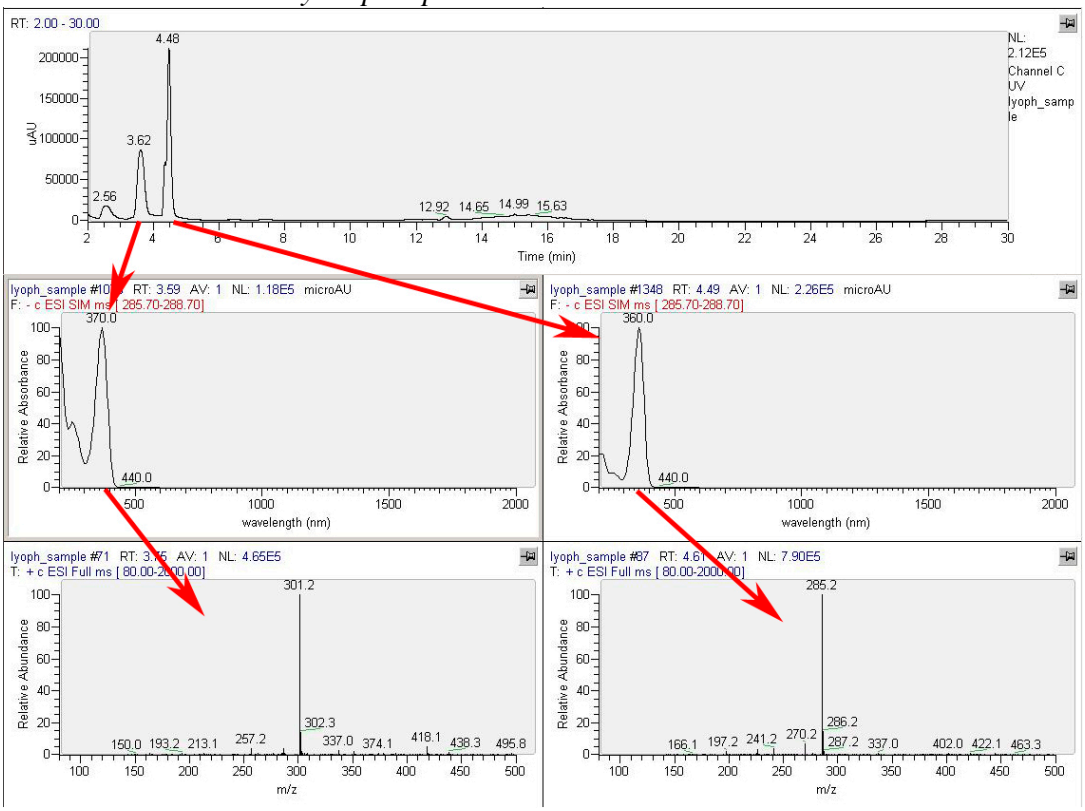


Figure 5

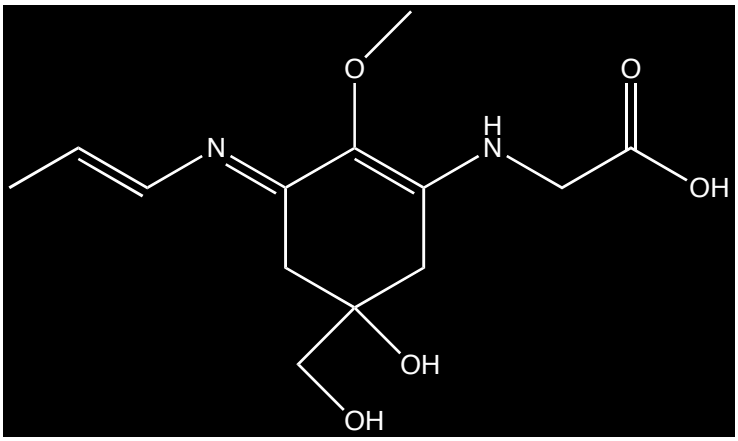
Major Accomplishment 5:

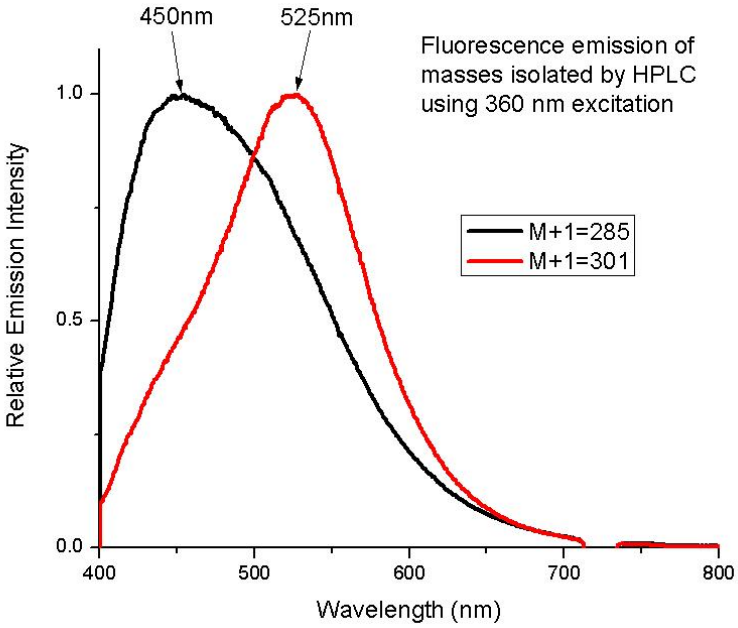
Bioluminescent tissues were dissected from *C. variopedatus* specimens, frozen in liquid N_2, and ground to a powder with a liquid N_2-chilled mortar and pestle. The powder was extracted with 50 mM Tris buffer, pH 7.2. A green fluorescent supernatant (360 nm exc.) was obtained.

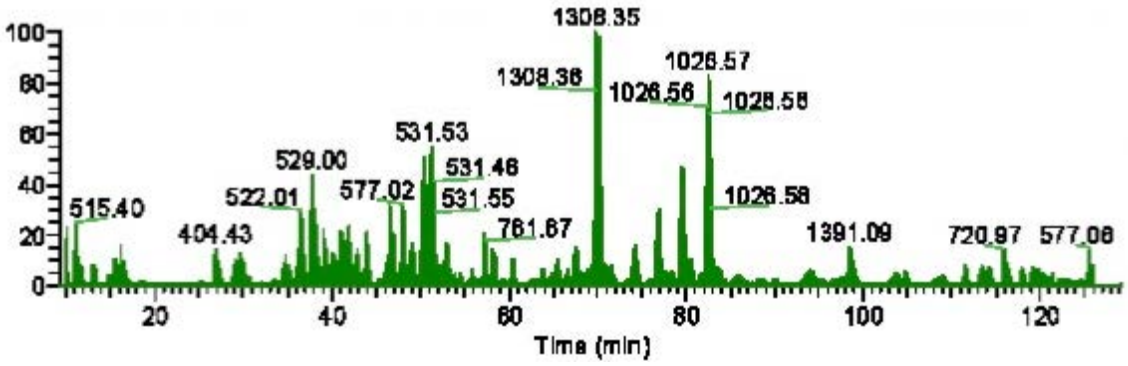
The aqueous extract of *C. variopedatus* run on LC/MS revealed two peaks with long wavelength abosrbance, 360 and 370 nm, corresponding with masses $(M+H)^+$ of 285 and 301, respectively. The former (285) was previously observed in a mucous extract of another bioluminescent worm *Odontosyllis phosphorea*.



Excitation of isolated peaks containing masses 284 and 300 at 360 nm in pH 8 buffer showed fluorescence emission peaks at 450 nm and 525 nm, respectively. HRMS revealed molecular formulae that differed by a single O atom. M = 284: $C_{12}H_{18}N_2O_5$ (λmax = 360 nm, 460 nm emission)and M = 300: $C_{12}H_{18}N_2O_6$ (λmax = 370 nm, 525 nm emission). The UV/Vis, hrms, and MS/MS data are consistent with M=284 being palythene, a mycosporine-like amino acid (MAA). The mass 300 is a putative novel MAA.







A gel band of mass ~ 20 kDa from worm extract was digested with trypsin and the tryptic peptides were analyzed by LC-MS/MS (Pfizer & Yale/Keck). Ferritin peptides were detected at

low levels (~2%). Results of separate proteomic and manual MS/MS sequencing-BLAST searches reveal a novel protein with mass 18,427 Da.

Chaetopterus variopedatus is one of the few known species of bioluminescent marine annelids. The tubeworm spends its adult-life entirely in an opaque parchment-like tube and filters food from the water column using a secreted mucous net (1, 2). Unlike most other polychaetes, C. variopedatus is capable of bright bioluminescence, with individuals able to secrete a luminescent slime upon neuro-stimulation (3-5). This behavior, however, is not definitively associated with mating, defense, or predation, and the purpose of this adaptation remains unclear (6). Likewise there is a relative lack of data concerning the biochemical nature of the worm's bioluminescent system beyond a few early studies in the 1960s (7, 8). Working with bioluminescent material harvested from thousands of specimens, Shimomura and Johnson determined that the light emission (λ_{max} ~460 nm) from worm extracts could be enhanced by Fe^{2+} plus peroxides. Shimomura hypothesized (9) that the bioluminescence system involved a photoprotein and possibly cofactors, but was unable to identify the structures of substrates or products of the light emitting reaction.

Our interest in the biochemical characterization of new bioluminescence systems led us to an investigation of the small molecule (non protein) components of the light reaction of this intriguing organism. Our strategy was to characterize the fluorescent components of worm extracts in an attempt to identify an intact or degraded substrate (luciferin) and/or emitter (oxyluciferin). In this report, we describe the results of our spectral and high performance liquid chromatography-electrospray ionization mass spectrometry (LC/ESIMS)-based studies and offer a working hypothesis for the marine worm bioluminescence process.

MATERIALS AND METHODS

Specimen collection: All worms used in this study originated from the La Jolla submarine canyon located in San Diego, CA. *Chaetopterus* lives in U-shaped tubes partially buried in soft seafloor sediment leaving only the end of the tubes exposed. In 2010-2013, tubes were uncovered and hand collected by SCUBA at 20-30 m depth, transported in collecting bags to a surface boat and then to the Marine Biology Experimental Aquarium Facility at Scripps Institution of Oceanography. There, they were kept with circulating seawater at ambient temperature until shipped overnight to Connecticut College where they were held in closed circuit aquariums with cold (15 ˚C) artificial seawater (Instant Ocean).

Preparation of worm extracts: Live C. variopedatus specimens were carefully extricated from their parchment tubes and placed in glass petri dishes wetted with artificial seawater. In a dark room, worm bioluminescence was stimulated by repeated addition of 0.3 – 0.5 mL aliquots of 0.1 M sodium phosphate buffer, pH 7.5 containing 0.5 – 1.0 M KCl until bioluminescence could not be detected with the naked eye. The bioluminescent slime was collected and used immediately or flash frozen in liquid N_2 and stored at -80 °C. Additional bioluminescent material was collected by dissection of the 12^{th} worm segment as described by Shimomura (9). This material was the most intensely luminescent, but was also the most viscous (9, 10).

The worm samples were heated at 60 °C for 1 h and centrifuged at 12,000 rpm to yield a non-viscous yellow supernatant. Alternatively, raw worm samples were centrifuged at 6,000 rpm

in Sartorius Vivaspin 5000 MWCO spin columns and the non-viscous flow-through was retained for analysis. In all cases, the samples extracted from luminescent material fluoresced green (~525 nm) with long-wave (366 nm) UV excitation. As necessary, extracts were lyophilized and reconstituted in smaller volumes for analyses.

Spectroscopic methods: Bioluminescence and fluorescence measurements were performed using a Horiba Jobin-Yvon imaging spectrometer equipped with a Xenon lamp excitation source and liquid N_2 cooled CCD detector. Data were collected at 22 °C in a 0.8 mL quartz cuvette with 5 nm slit widths and excitation at 375 nm for fluorescence measurements. Bioluminescence measurements were performed immediately on freshly harvested 12^{th} segment worm material. Typically, light emission decayed below detection within one min after collection. UV-Visible measurements were made using a Perkin Elmer Lambda 25 spectrophotometer.

Chromatography: Chromatographic analyses were performed using a ThermoFinnigan Surveyor LC system equipped with photodiode array detector. Flow leaving the PDA was split between a Finnigan Fluorescence Plus detector and a ThermoFinnigan LCQ Advantage of LTQ XL mass spectrometer with ESI probes. A 4.6x150 mm Phenomenex Luna C18 column was run at a flow rate of 0.6 mL/min using a water/acetonitrile gradient with 0.1% formic acid. The gradient increased from 5% to 30% acetonitrile over 25 min.

The typical qualitative approach to characterize a bioluminescence system is to prepare hot and cold extracts of light emitting organ material and to add the extracts back together to regenerate light. This result is interpreted as demonstrating that a luciferin substrate and any required cofactor(s) (in the hot extract) and a luciferase enzyme (active only in the cold extract) are present. While this approach works well for beetle luciferases and luciferins, it does not succeed with photoproteins like aequorin (9) because the substrate (coelenterazine) must be covalently bound to the apoprotein. In this study, we confirmed Shimomura's failure to produce light with this classical approach and concur that the marine worm's bioluminescence is likely produced by a photoprotein. We note too that, as previously reported (11,12), bioluminescence is associated with viscous material that tends to produce aggregates making it difficult to reproducibly perform characterization experiments. While we were aware that there are examples (9) of non-fluorescent and/or unstable luciferins and oxyluciferins, our initial approach was to prepare extracts of fluorescent worm material, and to separate, isolate and characterize the long wavelength (~300 nm to ~400 nm) absorbing materials. Substrates or products in a bioluminescence system producing emission at ~ 460 nm should have this absorption characteristic. Using LC/ESIMS with UV-Visible and fluorescence detection, we identified several chlorophyll-*a* metabolites, including pheophorbide-*a* (loss of phytol tail) and pheophytin-*a* (demetallated); several mycosporine-like amino acids (MAAs). including palythene (13); and tryptophan. The chlorophyll derivatives are suggestive of krill and dinoflagellate luciferin structures (14) and the possibility that the worm substrate is related has to be considered. The non-fluorescent MAAs and tryptophan are unlikely to be involved in the bioluminescence process.

We had previously measured the bioluminescence emission (λ_{max} = 463 nm) of the mucus produced by live worms (Figure 1a) and noted that the fluorescence emission spectrum displayed peaks at 461 nm and 519 nm (Figure 1b) shortly after the cessation of light emission. Because it is a very good match to the bioluminescence spectrum, the 461 nm fluorescence is likely from

the protein-bound oxyluciferin; however, the 519 nm component was initially puzzling. Moreover, on standing at room temperature for ~2 h, the shorter wavelength fluorescence peak nearly disappeared and a single peak with λ_{max} = 525 nm was present (Figure 1c). Since the fluorescence quantum yield of the oxyluciferin is unknown, it is not possible to determine whether the oxyluciferin has been transformed into the long wavelength emitting species or into a non-fluorescent product, revealing the long wavelength fluorescent compound. While imaging of the worm shows the green fluorescence concentrated in bioluminescent organs (Figure 2), there is no direct evidence relating the green fluorescence to bioluminescence. Nonetheless, we separated and isolated the source of the 525 nm fluorescence from mucus and identified it as riboflavin by comparison of its molecular weight, chromatographic and spectral properties to an authentic sample (Figure 3). While we occasionally detected FMN (riboflavin-5′-phosphate), we consistently (> 6 specimen collections) found that riboflavin was the major fluorescent compound in all extracts of luminescent material including the glowing slime. This finding is significant because riboflavin must be acquired from the environment.

While we recognize that the presence of riboflavin may simply be coincidental, we present, as a working hypothesis, that light emission in C. variopedatus is produced by a photoprotein version of bacterial bioluminescence. The direct involvement of bioluminescent bacteria was ruled out because samples of glowing mucus were plated (15) and failed to show any growth of luminescent colonies. However, we caution that in one instance we detected samples of the worms with exterior surface contamination by *Vibrio Campbellii*. a parasitic bioluminescent bacteria that emits at ~490 nm (16). We propose that riboflavin serves as the source of reduced riboflavin or $FMNH_2$, which is the non-fluorescent luciferin that is bound to an apoprotein forming a worm photoprotein, as is analogously the case with coelenterazine and the aequorin photoprotein. The occurrence of flavins has well-established precedence in bacterial bioluminescence (17, 18) in which $FMNH_2$ is a substrate produced by the action of a specific reductase on FMN. In a reaction also requiring O_2 and a long chain aldehyde, bacterial luciferase catalyzes light production (λ_{max} ~490 nm) from a 4-hydroxyflavin excited state intermediate (19) that is then further oxidized to FMN accompanied by light emission. The approximate range of bacterial bioluminescence emission is 472 nm to 500 nm (9). We believe it is feasible that reduced riboflavin produced by a reductase is "charged" to an apoprotein and that light emission is subsequently triggered. The emitter is likely a 4-hydroxyflavin sequestered by the apoprotein providing an environment that accounts for the reasonable stability of the emitter and its ability to produce 463 nm bioluminescence before it is further oxidized to riboflavin.

It is noteworthy that riboflavin has been proposed as a possible emitter of the hydrogen peroxide stimulated bioluminescence (λ_{max} = 528 nm) (20) of the luminous acorn worm Ptychodera flava. Bromohydroquinones and bromoquinones were isolated from this worm and a model study demonstrated that chemiluminescence (λ_{max} = 521 nm) could be produced from mixtures of riboflavin, hydrogen peroxide and 2,3,5,6- tetrabromohydroquinone (20). We believe that this is a less likely explanation for C. variopedatus bioluminescence because it would require a ~65 nm blue-shifted emission from riboflavin and we have been unable to identify these bromo compounds in ethyl acetate extracts of worm material. In addition, hydrogen peroxide in Chaetopterus appears to have an inhibitory effect on the light production from untreated luminous slime (9). Studies are in progress to identify the key proteins in luminous C. variopedatus sources to further characterize this fascinating bioluminescence process and to evaluate the working hypothesis presented here.

Figure 1.

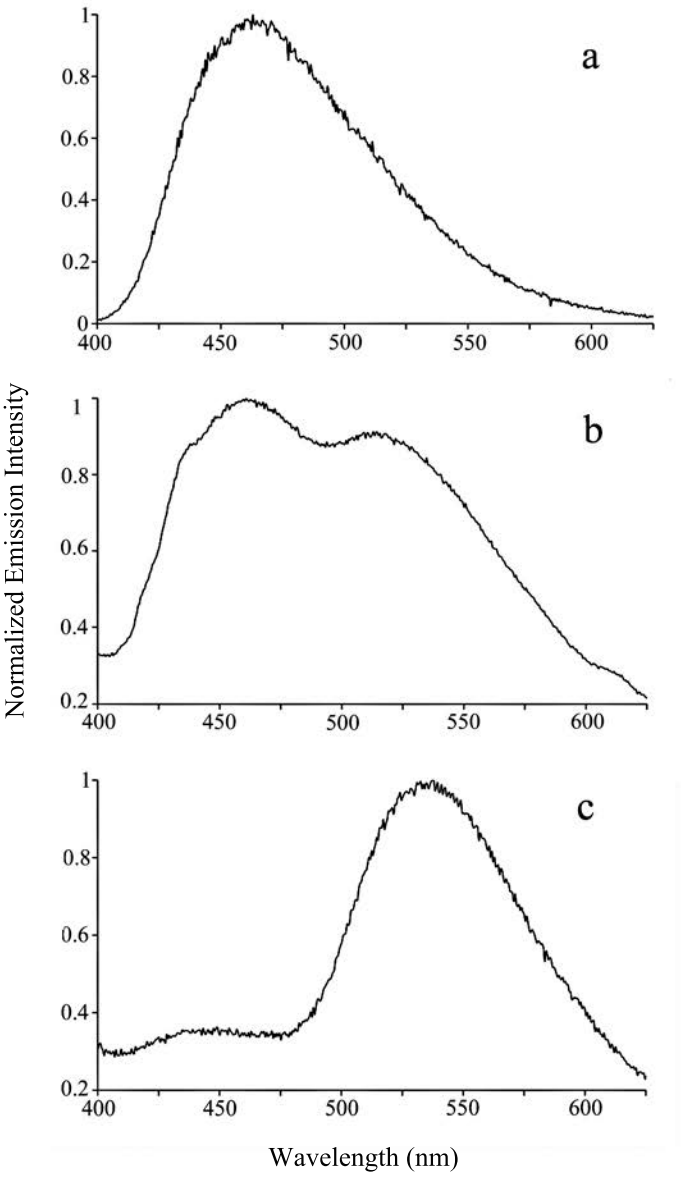


Figure 2.

Figure 3.

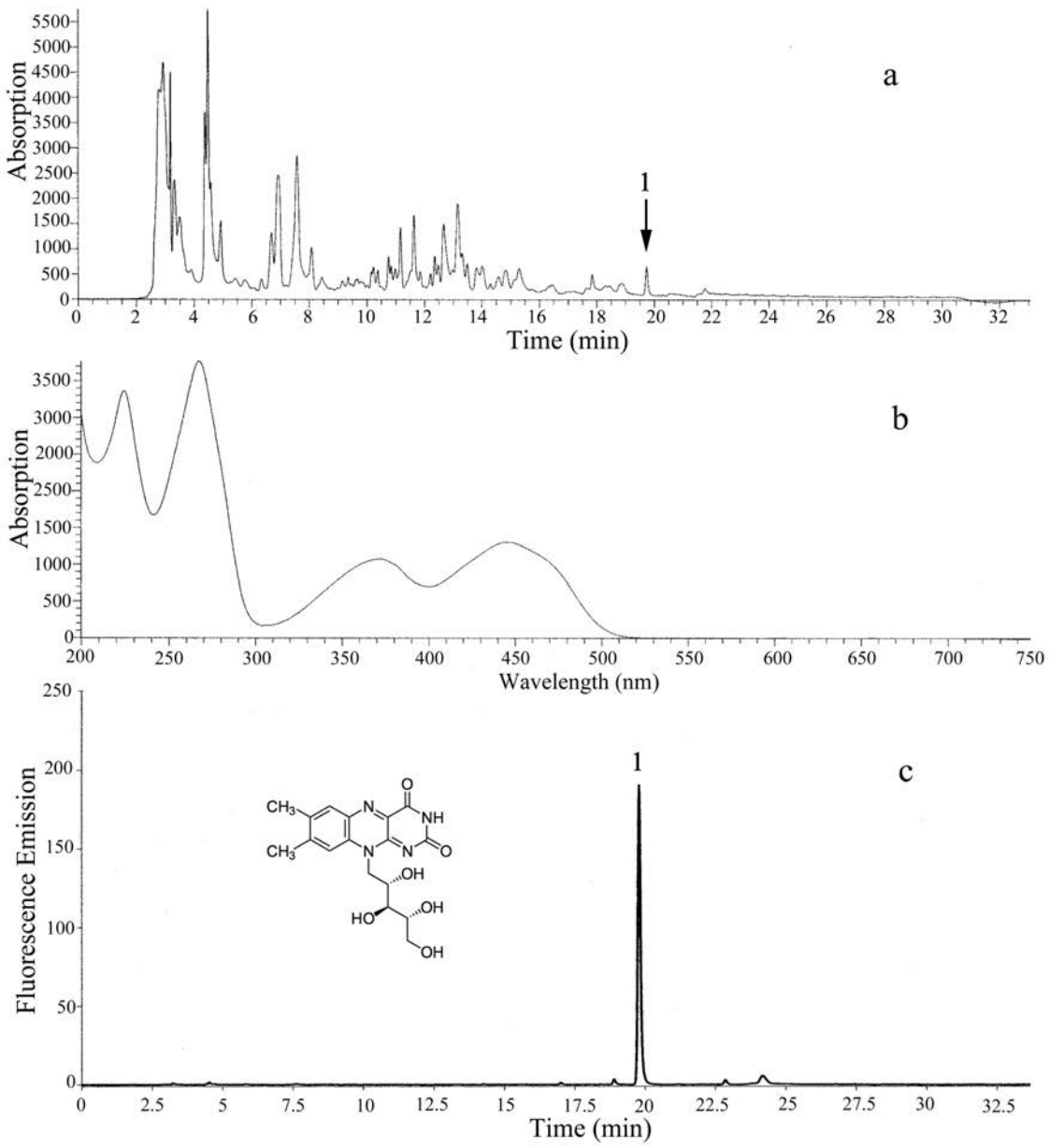

a
Absorption
Time (min)
1
b
Absorption
Wavelength (nm)
c
Fluorescence Emission
Time (min)
1
CH3
CH3
N
NH
O
OH
HO

FIGURE LEGENDS

Figure 1: Normalized *C. variopedatus* emission spectra. (a) Bioluminescence of mucus harvested from 12th segments of worms. (b) Fluorescence emission (375 nm excitation) of same sample immediately after cessation of detectable bioluminescence. (c) Fluorescence emission of spent luminescent material after standing at room temp for ~ 2 h.

2: Images in bright field and fluorescence (390 nm excitation) of the marine worm Chaetopterus.

Figure 3: Detection of riboflavin in extracts of luminescent *C. variopedatus* mucus. (a) HPLC chromatogram (PDA detection) showing the presence of riboflavin (peak 1) in worm extract. The mass of 377.2 ($M+H^+$) detected in this peak is consistent with that of riboflavin (M = 376.4). (b) UV-Visible spectra of peak 1, consistent with riboflavin standard. (c) Fluorescence chromatogram (375 nm excitation, 525 nm emission) of worm extract, showing riboflavin as major fluorescent component. Spiking samples with authentic riboflavin confirmed the identity of peak 1 as riboflavin.

REFERENCES

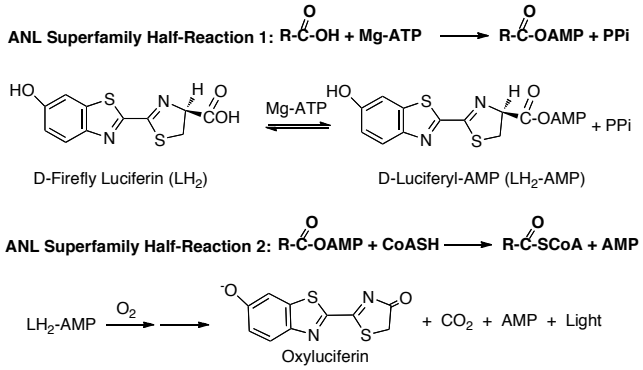
1. Flood, P. R. and A. Fiala-Médioni (1982) Structure of the mucous feeding filter of *Chaetopterus variopedatus* (Polychaeta). *Mar. Biol,* 72, 27-33.
2. MacGinitie, G. E. (1939) The method of feeding in *Chaetopterus*. *Biol. Bull.* 77, 115-118
3. Anctil, M. (1981) Luminescence control in isolated notopods of the tube-worm *Chaetopterus variopedatus* - Effects of cholinergic and gabaergic drugs. *Comp. Biol and Phys.* 68, 187-194.
4. Martin, N. and M. Anctil (1984) Luminescence control in the tube-worm *Chaetopterus variopedatus* - Role of nerve cord and photogenic gland. *Biol. Bull.* 166, 583-593.
5. Nicol, J. A. C. (1952) Studies on *Chaetopterus variopedatus*. III. Factors affecting the light response. *J. Mar. Biol. Assoc. U.K.* 31, 113-144.
6. Harvey, E. N. (1952) *Bioluminescence*. Academic Press, Inc., New York.
7. Shimomura, O. and F. H. Johnson (1966) Partial purification and properties of the Chaetopterus luminescence system. In *Bioluminescence in Progress* (Edited by F. H. Johnson and Y. Haneda), pp. 495-521.Princeton University Press, Princeton, NJ.
8. Shimomura, O. and F. H. Johnson (1968) Chaetopterus photoprotein: crystallization and cofactor requirements for bioluminescence. *Science* 150, 1239-1240.
9. Shimomura, O. (2006) *Bioluminescence: Chemical Principles and Methods*. World Scientific Publishing Co., Singapore.
10. Enders, H. E. (1909) A study of the life-history and habits of *Chaetopterus variopedatus*, Renier et Claparede. *J. Morphol.* 20, 479-531.
11. Johnson, F. H. (1959) Kinetics of luminescence in *Chaetopterus* slime, and the influence of certain factors thereon. *J. Cell. Comp. Physiol.* 53, 259-277.
12. Nicol, J. A. C. (1952) Studies on *Chaetopterus variopedatus* (Renier). I. The light-producing glands. *J. Mar. Biol. Assoc. U.K.* 30, 417-431.

13. Whitehead, K. and J. I. Hedges (2003) Electrospray ionization tandem mass spectrometric and electron impact mass spectrometric characterization of mycosporine-like amino acids. *Rapid Commun. Mass Spectrom.* 17, 2133-2138.
14. Topalov, G. and Y. Kishi (2001) Chlorophyll catabolism leading to the skeleton of dinoflaggellate and krill luciferins: hypothesis and model Ssudies. *Angew. Chem.* 113, 4010-4012.
15. Ulitzur, S. and J. W. Hastings (1979) Control of Aldehyde Synthesis in the Luminous Bacterium Beneckea-Harveyi. *J. Bacteriol.* 137, 854-859.
16. Suadee, C., S. Nijvipakul, J. Svasti, B. Entsch, D. P. Ballou and P. Chaiyen (2007) Luciferase from *Vibrio campbellii* is more thermostbale and binds reduced FMN better than its homologues. *J. Biochem.* 142, 539-552.
17. Hastings, J. W., C. J. Potrikus, S. C. Gupta, M. Kurfurst and J. C. Makemson (1985) Biochemistry and physiology of bioluminescent bacteria. *Adv. Microb. Physiol.* 26, 235-291.
18. Tu, S. C. and H. I. Mager (1995) Biochemistry of bacterial bioluminescence. *Photochem. Photobiol.* 62, 615-624.
19. Kurfuerst, M., P. Macheroux, S. Ghisla, and J. W. Hastings (1987) Isolation and Characterization of the Transient, Luciferase-Bound Flavin-4a-Hydroxide in the Bacterial Luciferase Reaction. *Biochim. Biophys. Acta* 924, 104-110.
20. Kanakubo, A. and M. Isobe (2005) Isolation of brominated quinones showing chemiluminescence activity from luminous acorn worm, Ptychodera flava. *Biorg. Med. Chem.* 13, 2741-2747.

Major Accomplishment 6:

The relationships between the conformational dynamics and functions of a large group of proteins termed the "ANL superfamily of adenylating enzymes," are the subject of an excellent recent review by Gulick[1]. The ANL superfamily enzymes that were first recognized by Conti and Brick[2] share ~20% sequence identity, are structurally similar, and include: the acyl-CoA synthetases, the adenylation domains of the non-ribosomal peptide synthetases (NRPS)[3], and the (beetle) luciferases. All ANL enzymes catalyze two partial reactions (Scheme 1), first converting carboxylate-containing substrates into the corresponding adenylates. With the exception of the luciferases, the second half-reaction involves the substitution of AMP by Coenzyme A (CoA) or protein-bound 4′-phosphopantethine producing a thioester. The 2nd partial reaction catalyzed by luciferase is quite different; a multi-step oxidative process that produces bioluminescence (Scheme 1).[4,5] Interestingly, the luciferase can use CoA, a cofactor not required for light emission, to convert dehydroluciferin-AMP (L-AMP), a potent inhibitor formed in an *in vitro* side reaction, into the less inhibitory thioester.[6]

Scheme 1. Partial reactions catalyzed by the ANL superfamily



Based on 4 crystal structures of ANL enzymes, only one of which contained CoA and a rotated C-domain, Gulick proposed[7] a domain alternation mechanism to account for the catalysis of the half-reactions (Figure 1). Substantial support for this mechanism for the acyl-CoA ligase subgroup is now available from sets of crystal structures of 2 enzymes[8,9] in both the adenylate- and thioester-forming conformations, which are related by ~140° rotations of the C-domain. However, the NRPSs and luciferases have only been crystallized in the adenylation conformation. This is especially problematic for luciferase biochemistry because: (1) CoA is not a required substrate; and (2) the oxidative half-reaction is mechanistically dissimilar to the thioester-forming reaction. However, we have reported[10,11] results from mutagenesis studies with *P. pyralis* luciferase (Ppy WT) that are consistent with a mechanism requiring 2 conformations. Specifically, we found that alanine replacement of Lys529 or Lys443, residues that are ~20 Å apart and on opposite faces of the C-domain, disrupted only the adenylation or oxidative partial reactions, respectively. With our results and crystal structures of luciferase in the adenylation conformation in hand,[2,12,13] we undertook this study to produce structural evidence for the existence of the rotated (thioester-forming) conformation of luciferase and to determine its relevance to the oxidative half-reaction that produces bioluminescence.

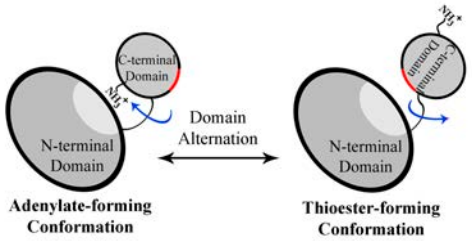


Figure 1. Schematic of domain alternation mechanism. The acyl-CoA synthetases and NRPSs in the ANL superfamily use an ~140° rotation to present opposite faces of the C-terminal domain to catalyze two different partial reactions.

Our experimental approach was designed to trap the elusive rotated luciferase conformation using Cys thiol chemistry. This required the introduction of Cys residues on each domain that could approach each other closely in the rotated conformation, but not in the adenylate-forming one. N-domain residue Ile108 and C-domain residue Tyr447 were initially chosen for mutation to Cys based on our observation that a crystal structure[14] of a highly thermostable long chain fatty acid CoA synthetase in the thioester conformation contained, among 35 salt bridges not found in Ppy WT, a single inter domain interaction involving Arg and Glu at the equivalent respective positions. Indeed, a luciferase variant containing the changes Ile108Arg and Tyr447Glu displayed 6-fold enhanced stability at 37 °C compared to Ppy WT. Because these residues are ~40 Å apart in the luciferase structures, it appeared that transient salt bridge formation had occurred in a rotated conformation. Next, the *P. pyralis* variant containing the changes Ile108Cys and Tyr447Cys was made. While this enzyme appeared to form intramolecular disulfide bonds under mild oxidizing conditions, the results were compromised by the apparent reactivity of the 4 native Cys residues. To circumvent this problem, we prepared the Ppy WT variant Ppy 9$^-$ in which the Cys residues were eliminated by mutation to Ala or Ser. To accomplish this, it was first necessary to make 5 additional amino acid changes[15] to overcome protein stability problems resulting in the formation of occlusion bodies encountered when variants containing multiple Cys changes were expressed. Ppy 9$^-$ produced strong bioluminescence with excellent specific activity based on total light emission, but displayed a 2-fold longer rise time and extended decay compared to Ppy WT (Table 1). Similar emission characteristics with more prolonged rise and decay times were observed with the addition of the Ile108Cys and Tyr447Cys changes that produced the protein Ppy 9$^-$ C108/C447.

While attempts to trap Ppy 9- C108/C447 in a rotated conformation through disulfide bond formation produced improved and encouraging results, we were unable to find suitable reaction conditions to cleanly accomplish the conversion. Instead, we adopted a chemical trapping approach using the symmetrical bifunctional reagent 1,2-bismaleimidoethane (BMOE) that we anticipated would function as shown in Figure 2 (and/or by conjugate addition first at Cys108 followed by

Table 1. Bioluminescence properties of luciferases at pH 7.8

Enzyme	Relative Activity[a]	Rise Time (s)[b]	Decay Time (min)[b]	λ_{max} [c]
Ppy WT	100 (100)	0.4	0.15	560
Ppy 9$^-$	37 (135)	0.8	2.90	562
Ppy 9$^-$ C108/C447	20 (119)	1.5	7.60	562

[a]Specific activity values based on flash heights and total light emitted (in parentheses). Activity assays contained LH_2 (100 μM) and Mg-ATP (2 mM) and are reported relative to Ppy WT values. [b]Time to reach maximum emission intensity and decay to 20% of this value. [c]Bioluminescence emission maximum.

reaction at Cys447). Upon incubation of a 20 μM solution of enzyme with 24 μM BMOE at 20 °C at pH 7.0, a rapid loss of bioluminescence activity was observed that was nearly complete (98%) in 1 h, following kinetics typical of irreversible enzyme inactivation . Exhaustive dialysis did not restore activity and LC/ESIMS analysis (Figure 3A) indicated a major component with mass 61 158 ±6, in agreement with the expected value (61 161) for BMOE covalently attached to Ppy 9$^-$ C108/C447 in a 1:1 molar ratio. Additionally, a minute peak (61 387 ±6) was detected corresponding to the covalent incorporation of 2 mol BMOE per mol of protein (Figure 3A). In a control experiment, treatment of BMOE-modified Ppy 9$^-$ C108/C447 with an excess of the sulfhydryl blocking reagent N-ethylmaleimide (NEM) produced no change in mass, indicating the absence of detectable free thiols. Taken together, the MS results are consistent with Ppy 9$^-$ C108/C447 being mostly intramolecularly cross-linked by BMOE, with a minor amount of the protein labeled twice presumably at Cys108 and Cys447. We further

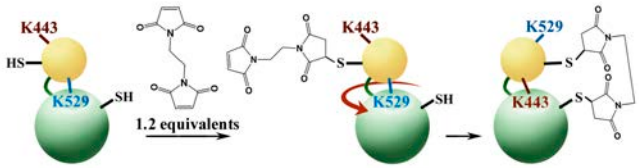


Figure 2. Chemical trapping of Ppy9$^-$ C108/C447 with BMOE. In the schematic, BMOE is shown attaching to C447 in the adenylate-forming conformation and, after C-domain rotation, to C108.

analyzed the BMOE-labeled luciferase using SDS-PAGE (Figure 3B). Ppy 9$^-$ gave one band corresponding to the expected mass of the enzyme. A band representing the same mass accounts for ~94% of the protein in a sample of Ppy 9$^-$ C108/C447, with the remainder eluting with an apparent mass of ~74 500. The higher mass band in the BMOE- labeled sample accounted for ~96% of the protein applied to the gel. Based on the MS data (Figure 3A), the higher mass band must represent the BMOE-modified luciferase. Apparently it migrates more slowly because it deviates from a typical globular shape and/or takes up an abnormally low amount of SDS consistent with our expectation of having trapped a single and different luciferase conformation. The minor band, corresponding to enzyme with 2 BMOE attached, would be expected to run normally because it cannot be locked into a single conformation. In the unmodified Ppy 9- C108/C447 sample, the slower moving minor band likely represents disulfide cross-linked enzyme, as we had previously observed. Although barely visible in Figure 3B, traces of protein linked intermolecularly by BMOE were also observed in the gel analysis.

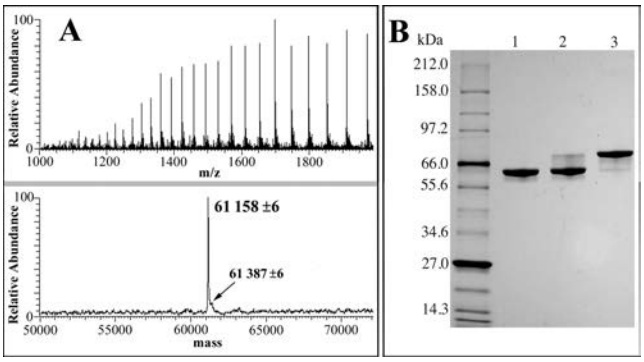


Figure 3. ESI/MS and SDS-PAGE analysis of Ppy 9- C108/C447 treated with BMOE. (A) The mass spectrum (top) and spectrum following BIOMASS deconvolution using the algorithm of the BioWorks 3.0 software (bottom) are shown. The 61 158 ±6 mass is in agreement with the expected mass of Ppy 9- C108/C447 (60 941) covalently modified with BMOE (220) in a 1:1 molar ratio and the minor peak (61 387 ±6) corresponds to the expected mass (61 381) for the protein labeled twice with BMOE. (B) SDS-PAGE gel (4%-20%) of Ppy 9-, Ppy 9- C108/C447 and BMOE-linked Ppy 9- C108/C447 in lanes 1-3, respectively.

Having demonstrated that Ppy 9- C108/C447 was covalently modified with BMOE, we used proteolysis and LC/ESIMS to confirm that the reagent had, in fact, cross-linked the protein through the position 108 and 447 Cys thiols as envisioned (Figure 2). Control experiments (**SI-5**) were first performed with unmodified and BMOE-treated Ppy 9- C108/C447 samples that were incubated with excess NEM. After digestion with chymotrypsin, the expected masses for the 2 NEM-containing peptides 98IGVAVAPANDCY109 and 445KGCQVAPAEL454 were readily detected in the unmodified sample, but not in the BMOE-treated one. Another BMOE-modified Ppy 9-C108/C447 sample was then sequentially digested with chymotrypsin, thermolysin and elastase (Figure 4). The masses 2425.9, 2123.0 and 1543.0 were detected after the protease treatments and correspond to those expected (2425.7, 2122.8 and 1542.6) from sequential cleavage at the sites indicated in Figure 4 by the red, blue and green arrows, respectively. We note that the calculated masses subsequent to the thermolysin treatment at 65 °C and pH 8.0 are based on the hydrolysis of both succinimide rings (+36) as verified in a control reaction. HRMS analysis of the final digest confirmed the molecular formula of a unique peptide (Figure 4) that confirms the cross-linking of Cys108 and Cys447. Since there was only trace evidence for intermolecular protein cross-linking and because BMOE can only span ~8 Å, it is not possible that the cross-linking occurred in the adenylate-forming conformation captured in the luciferase crystal structures in which residues 108 and 447 are ~40 Å apart. We conclude that luciferase has been covalently cross-linked by BMOE as shown in Figure 2, thus providing strong supporting evidence for the existence of a rotated luciferase conformation predicted by the domain alternation mechanism.

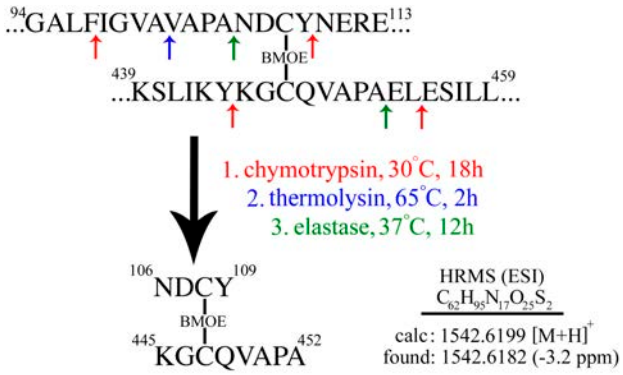


Figure 4. Proteolytic sequencing of Ppy 9$^-$ C108/C447 covalently labeled with BMOE. A 0.8 mg sample of the chemically modified protein was sequentially digested with immobilized chymotrypsin, thermolysin and elastase. After each digestion, masses were detected by LC/ESIMS corresponding to the calculated values for BMOE cross-linked peptides cleaved at the sites indicated by the colored arrows.

We next carried out light emission activity measurements (Figure 5) to investigate the role of the rotated conformation in luciferase bioluminescence. With the natural substrates LH_2 and Mg-ATP, light is emitted as a result of an overall process that requires both partial reactions (Scheme 1). The bioluminescence activities of Ppy 9$^-$ and Ppy 9$^-$ C108/C447 (Figure 5A) reflect their flash-height based specific activities (Table 1). After cross-linking with BMOE, however, Ppy 9$^-$ C108/C447 is essentially incapable of light production. The remaining activity can likely be ascribed to the small amount of enzyme labeled twice with BMOE. Luciferase bioluminescence also can be initiated by replacing the natural substrates with synthetic LH_2-AMP, thereby bypassing the adenylation partial reaction and requiring only that the oxidative 2nd half-reaction be functional. The relative light production from Ppy 9$^-$ and Ppy 9$^-$ C108/C447 produced with the synthetic adenylate (Figure 5B) is similar to that observed with natural substrates. In marked contrast to the result with LH_2 and Mg-ATP, light intensity observed with Ppy 9$^-$ C108/C447 after cross-linking with BMOE was restored and approached the level obtained with Ppy 9$^-$.

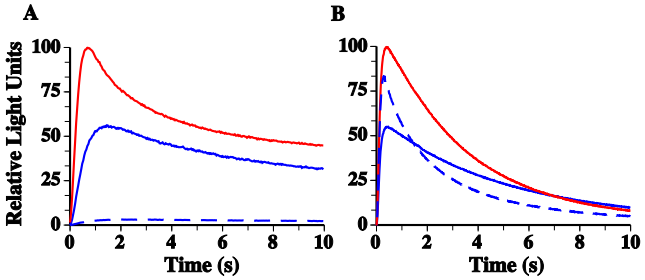


Figure 5. Bioluminescence time courses with natural substrates or synthetic LH_2-AMP. In addition to 0.6 μg protein (Ppy 9$^-$, solid red line; Ppy 9$^-$ C108/C447, solid blue line; or Ppy 9$^-$ C108/C447 cross-linked with BMOE, dashed blue line), 0.5 mL reaction mixtures in 50 mM glygly buffer, pH 7.8 contained (A) 100 μM LH_2 and 2 mM Mg-ATP or (B) 75 μM LH_2-AMP. Light emission was initiated by injection of Mg-ATP or enzyme.

Additional supporting evidence for the importance of the luciferase rotated conformation to oxidation, but not adenylation, was accumulated by estimation of the relative rates of the partial

reactions (Table 2). The introduction of Cys residues to Ppy 9$^-$ at positions 108 and 447 reduced the overall reaction rate 2-fold as a result of a similar drop in the oxidative half-reaction. BMOE cross-linking nearly restored the LH_2-AMP oxidation rate to the level of Ppy 9$^-$, evidently eliminating the deleterious effect of the Cys substitutions on the oxidation rate. The greatly reduced overall rate of light production (Figure 5A and Table 2) of the BMOE-trapped enzyme is predominantly due to an ~100-fold reduced rate of adenylate formation, as expected for a luciferase tapped in the oxidative conformation. The bioluminescence results strongly suggest that the trapped luciferase conformation closely, but not exactly, resembles the conformation populated during the course of the natural bioluminescence process. Another indication that the trapped conformation is perturbed is the red-shifted (612 nm) bioluminescence observed with the LH_2-AMP initiated reaction.

The previously documented[11] importance of Lys443 to only the oxidative chemistry was confirmed by introducing the Lys443Ala change into Ppy 9$^-$C108/C447 (Table 2). Reassuringly, this variant is nearly fully capable of catalyzing the adenylation half-reaction, but incapable of oxidizing LH_2-AMP. After BMOE cross-linking, however, the Lys443Ala–containing protein displayed a greatly reduced adenylation rate, while remaining incapable of producing light from LH_2-AMP. These results confirm the importance of the highly conserved luciferase β-hairpin motif ^{442}IKYKGYQV^{449} to only the oxidative luciferase reactivity. Interestingly, this motif is part of the pantetheine channel in the acyl-CoA thioester conformation[9].

Table 2. Relative rates of adenylation and oxidation half-reactions[a]

			Half-reactions	
Ppy 9$^-$ variant[b]	BMOE[c]	Overall	1st	2nd
Ppy 9$^-$	-	100	100	100
I108C/Y447C	-	56	113	52
I108C/Y447C	+	1.7	< 1.0	84
I108C/Y447C /K443A	-	0.5	93	0.6
I108C/Y447C /K443A	+	0.02	<1.0	0.4

[a] Relative overall and oxidation (2nd half-reaction) rates were estimated based on flash height-based activity values obtained with LH_2 and Mg-ATP or LH_2-AMP, respectively; the error associated with the rates is ±10%. Adenylation (1st half-reaction) rates were estimated from rates of L-AMP formation and the error associated with the rates is ±15%. Additional experimental details in Supporting Information. [b] Amino acid changes to Ppy 9$^-$ are indicated. [c]Enzyme treated with 1.2 equivalents of BMOE for 1 h at 20 °C.

We have demonstrated that a firefly luciferase variant containing Cys residues at positions 108 and 447 can be intramolecularly cross-linked by BMOE, trapping the enzyme in a C-domain rotated conformation previously undocumented in the available luciferase crystal structures[2,12,13]. The BMOE-modified protein cannot adenylate luciferin, but is nearly fully capable of bioluminescence as a result of its retained ability to carry out the half-reaction in which LH_2-AMP is converted into electronically excited state oxyluciferin. The cross-linked luciferase has apparently been trapped in a conformation similar to those adopted by acyl-CoA synthetases during the process of converting acyl-adenylates into the corresponding CoA thioesters.[8,9]

REFERENCES

(1) Gulick, A. M. *Acs Chemical Biology* **2009**, *4*, 811.
(2) Conti, E.; Franks, N. P.; Brick, P. *Structure* **1996**, *4*, 287.

(3) Strieker, M.; Tanovic, A.; Marahiel, M. A. *Current Opinion in Structural Biology* **2010**, *20*, 234.
(4) White, E. H.; Rapaport, E.; Seliger, H. H.; Hopkins, T. A. *Bioorg. Chem.* **1971**, *1*, 92.
(5) DeLuca, M. *Adv. Enzymol.* **1976**, *44*, 37.
(6) Fraga, H.; Fernandes, D.; Fontes, R.; Esteves da Silva, J. C. G. *FEBS Journal* **2005**, *272*, 5206.
(7) Gulick, A. M.; Starai, V. J.; Horswill, A. R.; Homick, K. M.; Escalante-Semerena, J. C. *Biochemistry* **2003**, *42*, 2866.
(8) Reger, A. S.; Wu, R.; Dunaway-Mariano, D.; Gulick, A. M. *Biochemistry* **2008**, *47*, 8016.
(9) Kochan, G.; Pilka, E. S.; von Delft, F.; Oppermann, U.; Yue, W. W. *J. Mol. Biol.* **2009**, *388*, 997.
(10) Branchini, B. R.; Murtiashaw, M. H.; Magyar, R. A.; Anderson, S. M. *Biochemistry* **2000**, *39*, 5433.
(11) Branchini, B. R.; Southworth, T. L.; Murtiashaw, M. H.; Wilkinson, S. R.; Khattak, N. F.; Rosenberg, J. C.; Zimmer, M. *Biochemistry* **2005**, *44*, 1385.
(12) Franks, N. P.; Jenkins, A.; Conti, E.; Lieb, W. R.; Brick, P. *Biophys. J.* **1998**, *75*, 2205.
(13) Nakatsu, T.; Ichiyama, S.; Hiratake, J.; Saldanha, A.; Kobashi, N.; Sakata, K.; Kato, H. *Nature* **2006**, *440*, 372.
(14) Hisanaga, Y.; Ago, H.; Nakagawa, N.; Hamada, K.; Ida, K.; Yamamoto, M.; Hori, T.; Arii, Y.; Sugahara, M.; Kuramitsu, S.; Yokoyama, S.; Miyano, M. *J. Biol. Chem.* **2004**, *279*, 31717.
(15) Branchini, B. R.; Ablamsky, D. M.; Murtiashaw, M. H.; Uzasci, L.; Fraga, H.; Southworth, T. L. *Anal. Biochem.* **2007**, *361*, 253.

Experimental Procedures

Materials

The following materials were obtained from the sources indicated: Mg-ATP (bacterial source), porcine elastase and thermolysin from Sigma-Aldrich (St. Louis, MO); 1, 2-bis-maleimidoethane (BMOE) from TCI America (Portland, OR); immobilized chymotrypsin from ProteoChem (Denver, CO); restriction endonucleases from New England Biolabs (Beverly, MA); mutagenic oligonucleotides from Integrated DNA technologies (Coralville, IA); and pGEX-6P-2 expression vector from GE Healthcare (Piscataway, NJ). Firefly luciferin (LH_2) and dehydroluciferyl-AMP (L-AMP) were generous gifts from Promega (Madison, WI) and Joaquim C.G. Esteves da Silva. The found molecular masses (Da) of the following new proteins were within the allowable experimental error (0.01%) of the calculated values (in parenthesis): Ppy I108R, 61

202 (61 200); Ppy Y447E, 61 122 (61 123); Ppy WT I108R/Y447E, 61 166 (61 166); Ppy 9$^-$, 61 008 (61 012); Ppy 9$^-$ C108/C447, 60 938 (60 941) and Ppy 9$^-$ C108/C447/K443A, 60 884 (60 885).

General methods

Protein concentrations were determined with the Bio-Rad Protein Assay system using BSA as the standard. Site-directed mutagenesis was preformed with the QuikChange® Lightning Site-Directed Mutagenesis kit from Stratagene (La Jolla, CA). DNA sequencing to verify all mutations was performed at the W. M. Keck Biotechnology Laboratory at Yale University. Specific activity measurements were determined as previously reported [1,2] except that the integration time was 15 min and the final LH_2 concentration was 0.1 mM. Bioluminescence emission spectra were obtained using methods and equipment previously described.[3] The conditions used for reactions in which LH_2-AMP was the sole substrate are given in the legend to Figure S3. The relative overall and oxidative half-reaction rates, using flash heights as estimates of initial velocities, were estimated as described previously[4,5] except that final concentrations of 0.1 mM LH_2 and 75 μM LH_2-AMP were used. Also, integration-based activities were recorded until at least 90% of the total light was collected. The relative adenylation rates were estimated by fluorescence-based assays of dehydroluciferin-AMP formation according to prior reported methods.[4,5] Ppy WT and the new luciferases used in this study were expressed as GST-fusion proteins and purified by affinity chromatography (yields of ~5mg/0.25 L culture) and stored as described in detail in a previous report.[6] Initially, DTT was omitted from purifications of thiol-containing enzymes when oxidation and reaction with BMOE was planned. However, in some cases, disulfide formation was observed requiring treatment with TCEP (2 mM for 30 min at 22 °C) to restore the free thiols. Later, DTT was included in all purifications and removed by exhaustive dialysis prior to chemical cross-linking experiments. Heat inactivation studies to estimate thermostability were performed as described previously.[7] Potassium ferricyanide (FeCy) oxidation studies were performed by incubating proteins (20 μM) in 50 mM sodium phosphate buffer pH 7.0 containing 100 mM NaCl and 1 mM EDTA (MR buffer) at 4 °C with FeCy (0.4 mM or 50 μM) for 18 h. Mass spectral analyses were performed by tandem HPLC-electrospray ionization mass spectrometry (LC/ESIMS) using a ThermoFinnigan Surveyor HPLC system and a ThermoFinnigan LCQ Advantage mass spectrometer. The conditions for protein mass determinations were: column, Jupiter 5 μm C4 300Å (50 *x* 1.00 mm); wavelength, 270 nm; mobile phase, 95:5 water (0.1% TFA):acetonitrile (0.1%

TFA), gradient after 5 min to 5:95 water (0.1% TFA):acetonitrile (0.1%TFA) over 5 min; flow rate, 0.05 mL/min; MS mode, ES+; scan range, m/z = 200-2000; scan time, 0.2s. The electrospray source of the MS was operated with a capillary voltage of 37 V, and source voltage of 3.5 kV. Total mass spectra for protein samples were reconstructed from the ion series using Bioworks Browser 3.0 with BIOMASS deconvolution. Protease digests were analyzed with a Jupiter Proteo column (100 *x* 1.00 mm) and mobile phase, 95:5 water (0.1% TFA):acetonitrile (0.1% TFA), gradient after 5 min to 5:95 water (0.1% TFA):acetonitrile (0.1%TFA) over 45 min; flow rate, 0.15 mL/min; monitored at 220 nm.

Cloning of Ppy I108R, Ppy Y447E and Ppy I108R/Y447E

Starting with Ppy WT in the pGEX-6P-2 vector as the template,[8] the I108R and Y447E changes were introduced individually to make Ppy I108R and Ppy Y447E, respectively. The double mutant was made by introducing the Y447E into the plasmid containing Ppy I108R. The following primers and their respective reverse compliments were used: Y447E, 5´- CT TTA ATT AAA TAC AAA GGA **GAG** CAG GTG GCC CCC GCT G-3´ [*Eco*RV] and I108R, 5´- GGA GTT GCA GTG GCG CCC GCG AAC GAC **CGT** TAT AAT GAA CGT-3´ [*Kas*I] (bold represents the mutated codons, underlined represents the silent changes to create or remove a unique screening endonuclease site and brackets indicate the screening endonuclease).

Cloning of Ppy 9$^-$, Ppy 9$^-$ C108/C447 and Ppy 9$^-$ C108/C447/K443A

Starting with Ppy WT-TS (contains the amino acid changes T214A, A215L, I232A, F295L and E354K) in the pGEX-6P-2 vector,[6] Ppy 9$^-$ was made by mutating the endogenous Cys residues using the indicated primers and their respective reverse compliments: C81S, 5´- CAC AGA ATC GTC GTA **AGC** AGC GAA AAC TCT CTT CAA TTC -3´ [*Bts*I]; C216A, 5´- CAT AGA GCT CTC **GCC** GTC AGA TTC TCG CAT GCC AGA GAT CC-3´ [*Sph*I]; C258S, 5´- GA ATG TTT ACT ACA CTC GGA TAT CTG ATA **TCT** GGA TTT CGA GTC GTC -3´ [*Eco*RV]; C391S, 5´- CAG AGA GGC GAA TTA **TCT** GTC AGA GGG CCT ATG ATT ATG -3´ [*Ppu*MI]. Ppy 9$^-$ C108/C447 was made from the Ppy 9$^-$ plasmid by sequentially introducing the Y447C and I108C changes using the following primers and their respective reverse compliments: Y447C, 5´- G AAG TCT TTA ATA AAA TAC AAA GGA **TGT** CAG GTG GCC CCC GCT G -3´ [*Pac*I] and I108C, 5´- GGA GTT GCA GTG GCG CCC GCG AAC GAC **TGT** TAT AAT GAA CGT G -3´ [*Nar*I]. Ppy 9$^-$ C108/C447/K443A was constructed from the C108/C447 plasmid using the primer 5´- ATA GTT GAC CGC CTG AAG TCT TTA ATA **GCA** TAC AAA GGA TGT CAG GTG G-3´ [*Acu*I].

Modification of luciferases with 1, 2-Bismaleimidoethane (BMOE)

Solutions of Ppy 9$^-$, Ppy 9$^-$ I108C/Y447C or Ppy 9$^-$ I108C/Y447C/K443A (20 μM) in MR buffer were incubated with BMOE (24 μM) at 20 °C for 1 h. Aliquots (2 μL) were removed and assayed with 100 μM LH_2 and 2 mM Mg-ATP to determine the percent remaining bioluminescence activity. Incubation mixtures were exhaustively dialyzed against MR buffer and aliquots were analyzed by LC/ESIMS prior to and after treatment with 2 mM *N*-ethylmaleimide (NEM) for 1 h at 20 °C to determine free thiol content and the stoichiometry of BMOE incorporation.

Identification of Peptides Cross-Linked with BMOE

Unmodified and BMOE cross-linked Ppy 9$^-$ C108/C447 samples (0.8 mg in 0.6 mL MR buffer) were incubated with 2 mM NEM for 1 h at 20°C and dialyzed into 50 mM ammonium bicarbonate buffer, pH 8.0. The protein samples were added to 0.2 mL of a 50% slurry of immobilized chymotrypsin on agarose resin (ProteoChem) and incubated at 30 °C for 12 h with gentle mixing. The mixtures were centrifuged at 3 000 rpm for 3 min and the supernatants were retained. Analysis of aliquots (20 μL) by LC/ESIMS revealed chromatography peaks containing masses of 1317.4 and 1140.5, corresponding to the expected masses of NEM-containing peptides ^{98}IGVAVAPANDCY^{109} (1317.5) and ^{445}KGCQVAPAEL^{454} (1140.2), in the digest of the unmodified enzyme. The BMOE-treated protein digest did not contain these peptides, but did have a unique peak corresponding to the BMOE crossed link peptide shown in Figure 4.

The remainder of the BMOE-treated Ppy 9$^-$ C108/C447 chymotrypsin digest was incubated with 0.5 μM thermolysin at 65 °C for 2 h and the proteolysis was quenched with EDTA (final concentration of 1 mM). A 20 μL aliquot was removed for analysis and the remaining solution was incubated with 1.2 μM porcine elastase at 37 °C for 12 h. Aliquots of this solution and the chymotrypsin-thermolysin were analyzed by LC/ESIMS, which revealed masses corresponding to the cross-linked peptides shown in Figure 4. A sample (0.1 mL) of the triple protease digest was lyophilized and submitted for HRMS analysis by FT-ICR MS at the W. M. Keck Foundation Biotechnology Resource Laboratory at Yale University. The results of the analysis are presented in Figure 4 along with the sequence of the unique cross-linked peptide in which each maleimide ring was hydrolytically opened

II. TABLES. (All values in the tables below were obtained from at least triplicate trials and are reported as mean ± standard deviation.)

Table S1. Specific Activity and Thermostability of Luciferase Enzymes

Enzyme	Relative Specific Activity[a] Flash Height	Integrated	Thermal Inactivation (h)[b]	T_m (°C)[c]
Ppy WT	100 ± 2	100 ± 2	0.6 ± 0.06	43.9 ± 0.02
Ppy I108R	60 ± 2	110 ± 2	0.8 ± 0.1	45.5 ± 0.03
Ppy Y447E	44 ± 3	51 ± 2	0.9 ± 0.1	44.5 ± 0.02
Ppy I108R-Y447E	35 ± 2	87 ± 6	3.5 ± 0.3	46.4 ± 0.02

[a]Specific activities were obtained at pH 7.8 with LH_2 and Mg-ATP and are expressed relative to Ppy WT values, which are defined as 100. Integrated activity based on total light emission in 15 min. [b]Time to reach 50% remaining activity at 37°C as described[7]. [c]Mean aggregation temperature determined in 0.1 M sodium phosphate buffer (pH 7.5) containing 1 mM EDTA by CD spectroscopy as previously described[6].

Table S2. Chemical Modification Controls with Ppy WT

Reagent[a]	Thiols[b]	Relative Activity[c]
None	4.0 ± 0.1	100 ± 2
FeCy	1.2 ± 0.1	6[d] ± 0.5
BMOE	3.4 ± 0.2	66 ± 2

[a] Ppy WT (20 μM) samples were treated with 0.4 mM potassium ferricyanide (FeCy) for 18 h at 4 °C or with 24 μM BMOE for 1 h at 20 °C. [b]Determined by Ellman's assay[9]. [c]Specific activities were obtained at pH 7.8 with LH_2 and Mg-ATP and are expressed relative to Ppy WT values, which are defined as 100. [d]Value increased to 53 after treatment with 40 mM DTT for 18 h at 4 °C.

Table S3. Bioluminescence Properties of Luciferases at pH 7.8

Enzyme	Relative Activity[a]	Rise Time (s)[b]	Decay Time (min)[b]	λ_{max} [c] (nm)
Ppy WT	100±2	0.4±0.05	0.15±0.01	560±1
Ppy 9$^-$	135±1	0.8±0.03	2.90±0.18	562±1
Ppy 9$^-$ C108/C447	119±3	1.5±0.07	7.60±0.35	562±1

[a]Specific activity values based on the total light emitted. Activity assays contained LH_2 (100 μM) and Mg-ATP (2 mM) and are reported relative to Ppy WT values. [b]Time to reach maximum emission intensity and decay to 20% of this value. [c]Bioluminescence emission maximum.

Table S4. Mild Oxidation of Ppy 9$^-$ C108/C447 with FeCy

FeCy[a]	TCEP[b]	Thiols	Relative Activity[c]
-	-	2.0 ± 0.1[d]	100 ± 3
+	-	0.3 ± 0.04[d,e]	30 ± 1
+	+	2.0 ± 0.2[f]	78 ± 1

[a] Proteins were incubated with 50 μM FeCy for 18 h at 4 °C. [b]Treatment with 2 mM TCEP for 2 h at 4 °C. [c]Specific activities were obtained at pH 7.8 with LH_2 and Mg-ATP and are expressed relative to Ppy WT values, which are defined as 100. [d]Determined by Ellman's assay[9]. [e]Non-reducing SDS-PAGE contained bands for intramolecularly cross-linked enzyme (moderate), enzyme that was not cross-linked (weaker) and intermolecular cross-linked protein (trace). [f]Determined by NEM treatment and LC/ESIMS.

Figures

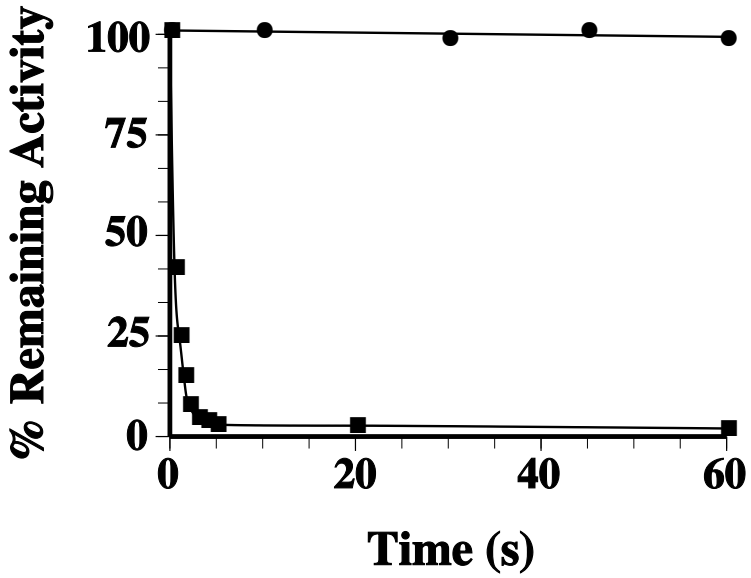


Figure S1. Inhibition of bioluminescence activity of Ppy 9$^-$ C108/C447 by BMOE. Samples of Ppy 9$^-$ C108/C447 (20 μM in MR buffer) were incubated in the absence (solid circles) or presence (solid squares) of 24 μM BMOE at 20 °C. The percent remaining bioluminescence activity values were determined from activity measurements performed by adding aliquots (2 μL) of incubation mixtures to 0.4 mL of 25 mM glycylglycine buffer, pH 7.8 containing 135 μM LH_2 in tubes placed into a custom luminometer[2]. Reactions were initiated by the rapid injection of 0.12 mL of 9.0 mM Mg-ATP in the same buffer. Additional experimental details including data acquisition and handling are available[1,2]. In separate experiments, untreated Ppy 9$^-$ and a sample incubated with BMOE gave essentially the same result as the control (solid circles).

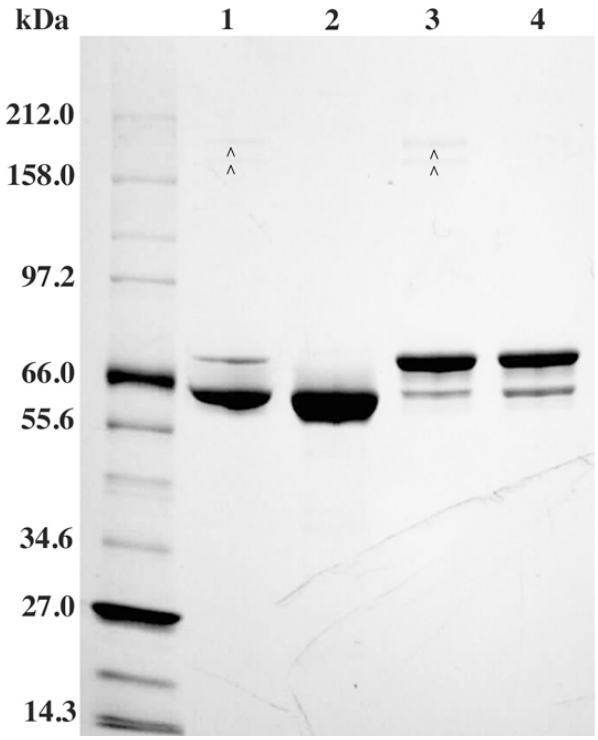


Figure S2. SDS-PAGE analysis of Ppy 9$^-$ C108/C447 treated with BMOE and TCEP. Non-reducing SDS-PAGE gel (4%-20%) of a Ppy 9$^-$ C108/C447 sample stored for 1 day at 4 °C without DTT prior to (lane 1) and after treatment with TCEP (lane 2). A Ppy 9$^-$ C108/C447 sample cross-linked with BMOE before (lane 3) and after treatment with TCEP (lane 4).

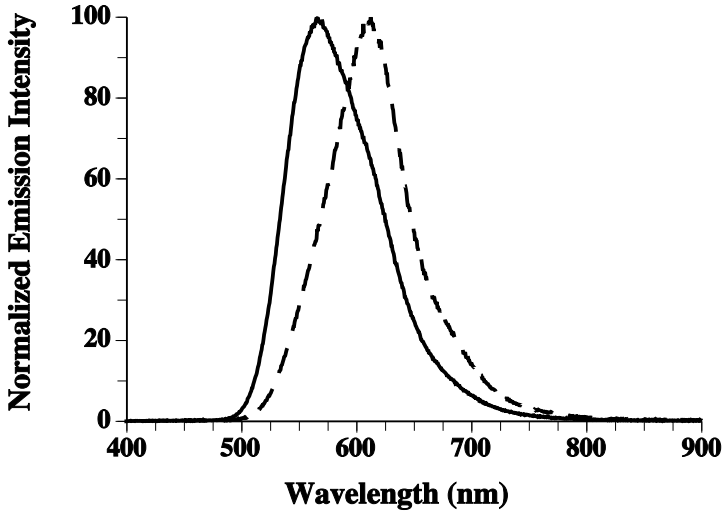


Figure S3. Bioluminescence emission spectra with synthetic LH_2-AMP. Reactions (0.51 mL) were initiated in 50 mM glycylglycine buffer, pH 7.8, containing 75 μM LH_2-AMP by the addition of 10 μL (0.6 μg in MR buffer) of Ppy 9$^-$ C108/C447 prior to (solid line) and after cross-linking with BMOE (dashed line).

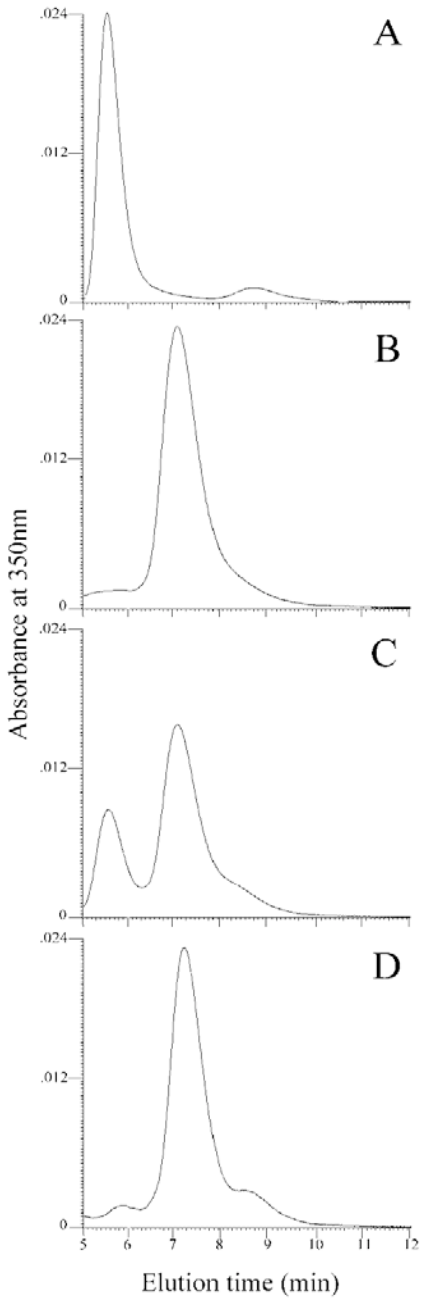


Figure S4. Luciferase catalyzed synthesis of L-CoA from L-AMP. The RP-HPLC methodology reported by Fraga *et al*.[10] was used to analyze mixtures of luciferases incubated with L-AMP. Solutions (0.1 mL) containing 20 μM L-AMP, 200 μM CoA, 4 mM $MgCl_2$, and 0.5 μM enzyme were incubated at 22 °C in Hepes buffer, pH 8.2 quenched by addition of an equal volume of MeOH:H_2O (2:1, v/v) and centrifuged for 1 min at 12,000 rpm. Supernatants (20 μL) were injected onto a Phenomenex Luna C18 3 μm column (100 mm x 1 mm) eluted with 2.9 mM sodium phosphate, pH 7.0: MeOH (2:1, v/v) at 75 μL/min. Total absorbance and absorbance at 350 nm were monitored and peak identities were verified by ESIMS. (**A**) control (no enzyme), only L-AMP is detected; (**B**) Ppy 9^- C108C447, 10 min incubation, only L-CoA detected; (**C**) BMOE cross-linked Ppy 9^- C108C447, 10 min, unreacted L-AMP and L-CoA detected; and (**D**) BMOE cross-linked Ppy 9^- C108C447, 30 min, only L-CoA detected.

References and Abbreviations

(1) Branchini, B. R.; Southworth, T. L.; Khattak, N. F.; Michelini, E.; Roda, A. *Anal. Biochem.* **2005**, *345*, 140.
(2) Branchini, B. R.; Ablamsky, D. M.; Rosenman, J. M.; Uzasci, L.; Southworth, T. L.; Zimmer, M. *Biochemistry* **2007**, *46*, 13847.
(3) Branchini, B. R.; Ablamsky, D. M.; Rosenberg, J. C. *Bioconjugate Chem.* **2010**, *21*, 2023.
(4) Branchini, B. R.; Murtiashaw, M. H.; Magyar, R. A.; Anderson, S. M. *Biochemistry* **2000**, *39*, 5433.
(5) Branchini, B. R.; Southworth, T. L.; Murtiashaw, M. H.; Wilkinson, S. R.; Khattak, N. F.; Rosenberg, J. C.; Zimmer, M. *Biochemistry* **2005**, *44*, 1385.
(6) Branchini, B. R.; Ablamsky, D. M.; Murtiashaw, M. H.; Uzasci, L.; Fraga, H.; Southworth, T. L. *Anal. Biochem.* **2007**, *361*, 253.
(7) Branchini, B. R.; Ablamsky, D. M.; Davis, A. L.; Southworth, T. L.; Butler, B.; Fan, F.; Jathoul, A. P.; Pule, M. A. *Anal. Biochem.* **2010**, *396*, 290.
(8) Branchini, B. R.; Magyar, R. A.; Murtiashaw, M. H.; Anderson, S. M.; Zimmer, M. *Biochemistry* **1998**, *37*, 15311.
(9) Ellman, G. L. *Arch. Biochem. Biophys.* **1959**, *82*, 70.
(10) Fraga, H.; Fernandes, D.; Fontes, R.; Esteves da Silva, J. C. G. *FEBS Journal* **2005**, *272*, 5206.

Abbreviations used: BMOE, 1, 2-bismaleimidoethane; FeCy, potassium ferricyanide; LC/ESIMS, tandem HPLC-electrospray ionization mass spectrometry; L, dehydroluciferin; L-AMP, dehydroluciferyl-AMP; L-CoA, CoA thioester of dehydroluciferin; LH_2, D-firefly luciferin; LH_2-AMP, luciferyl-AMP; MR buffer, 50 mM sodium phosphate buffer pH 7.0 containing 100 mM NaCl and 1 mM EDTA; NEM, *N*-ethylmaleimide; Ppy WT, recombinant *Photinus pyralis* luciferase containing the additional N-terminal peptide GPLGS; Ppy 9^-, Ppy WT containing the mutations C81S, T214A, A215L, C216A, C258S, C391S, I232A, F295L and E354K; and TCEP, tris(2-carboxyethyl)phosphine.

Major Accomplishment 7:

Firefly luciferase is a member of a large superfamily of adenylating enzymes. This ANL superfamily, named after the acyl-CoA synthetases, the adenylation domains of the modular non-ribosomal peptide synthetases (NRPSs), and luciferase, is present throughout all kingdoms of life and plays critical roles in both primary and secondary metabolism[1]. The three sub-families catalyze two-step reactions, sharing an initial adenylating step to produce an acyl-AMP intermediate. The adenylate then serves as a substrate for a second step that is specific to the particular subfamily. The acyl-CoA synthetases and NRPS adenylation domains both use the adenylate for a thioester-forming reaction, using the pantetheine thiol of either CoA or a holo-acyl carrier protein domain as an acceptor of the activated acyl group. In contrast to these reactions, luciferase catalyzes a multistep oxidative decarboxylation of the luciferyl-AMP intermediate (LH2-AMP) to produce bioluminescence.

Structures of ANL enzymes illustrate a 400-500 residue N-terminal domain and a smaller C-terminal domain of ~110-130 amino acids. The active site of these enzymes is located at the interface between these two domains. Ten conserved regions of these proteins have been termed the A1-A10 motifs[2]; sever-al of these motifs play critical roles in either or both partial reactions1. ANL enzymes are thought to use a domain alternation catalytic strategy in which the initial adenylation reaction is catalyzed by one conformation. Upon formation of the adenylate and release of pyrophosphate, a ~140° rotation of the C-terminal domain allows the enzymes to adopt a catalytic con-formation that is used for the second partial reaction. This catalytic strategy is supported by crystal structures of multiple acyl-CoA synthetases bound to CoA or CoA analogs[3-6] and, recently, of NRPS adenylation domains bound to the acyl-carrier proteins[7,8]. Mutations to residues on opposing faces of the C-terminal domain[9,10] and the hinge residue[7,9,11] have also confirmed this catalytic strategy.

For luciferase, however, only biochemical evidence sup-ports this catalytic strategy. Two catalytic lysine residues, one on each face of the C-terminal domain, are required for each partial reaction in a manner that suggests that a similar domain rotation is required for catalysis of the full reaction[12,13]. Mutation of Lys529, the A10 lysine, impairs only the adenylation reaction while mutation of Lys443 in the A8 region disrupts the oxidative reaction. Recently, an engineered luciferase enzyme containing cysteine substitutions was designed to test the domain movement hypothesis14. The two targeted residues are 37 Å apart in L. cruciata luciferase in the adenylate-forming conformation15 but should be susceptible to cross-linking if the enzyme can adopt the second conformation. Indeed, the luciferase variant could react with a chemical cross-linker. Importantly, this trapped enzyme was competent for catalysis of the oxidative reaction when provided with the luciferyl-adenylate intermediate (LH2-AMP) and the activity is dependent on the side chain of Lys443. We have successfully crystallized this trapped luciferase enzyme[14] and present here the first crystal structure of luciferase in this second catalytic conformation.

Structural determination of the chemically cross-linked luciferase enzyme. Ppy 9- C108/C447 is a P. pyralis luciferase variant in which the four native cysteine residues were eliminated by mutation to serine or alanine and two surface cysteine residues were introduced with the I108C and Y447C changes[14]. The structure of this enzyme, bound to the adenylate analog 5'-O-[N-

dehydroluciferyl)-sulfamoyl]adenosine (DLSA)16 and treated with the chemical cross-linker 1,2-bis(maleimido) ethane, was solved by molecular replacement using the N-terminal domain of luciferase[17]. The C-terminal domain was identified through iterative cycles of manual model building and refinement. The two chains of the asymmetric unit are very similar, with a root-mean-square displacement of all C α positions of 0.4 Å. For the purposes of comparison, we also solved a new crystal structure of the wild-type P. pyralis luciferase enzyme in the adenylate-forming conformation bound to DLSA (Figure 1A).

Comparison of luciferase enzymes in adenylate-forming and second conformations. The Ppy 9-C108/C447 luciferase enzyme crystallized in the conformation identified previously to be used by thioester-forming ANL enzymes to catalyze the second partial reaction (Figure 1B). Following the conserved A8 motif harboring the hinge residue at Lys439, the subsequent antiparallel two stranded β -sheet is directed into the active site of the enzyme. The ϕ / ϕ angles of Lys439 change from -73°/-12° in the structure of wild-type luciferase in the adenylate-forming conformation to -69°/158° in the cross-linked structure. As with other members of the ANL family of enzymes, this illustrates that a large component of the conformational change occurs with a rotation of the ϕ angle of the hinge residue. Additional torsion angle changes are seen in ϕ angles the Arg437 and Leu441, although the magnitude of the change is not as large as at the hinge residue Lys439.

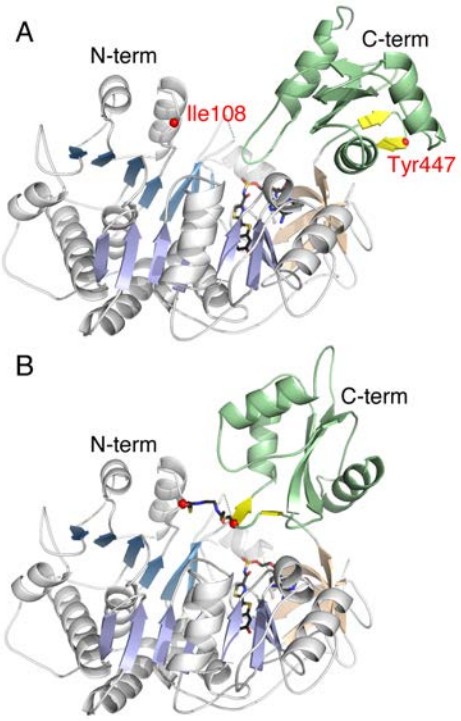


Figure 1. Structure of luciferase in both conformations. A. The wild-type luciferase in the adenylate-forming conformation bound to DLSA. The location of Ile108 and Tyr447 are highlighted in red. B. The cross-linked luciferase crystallized in the second conformation

The active sites of the two luciferase structures illustrate conserved interactions that have been seen in other adenylate-forming enzymes of the ANL superfamily. In the adenylate-forming conformation, Lys529, the catalytic lysine for the initial adenylation reaction interacts with the carbonyl oxygen of the adenylate, the O5 atom that bridges the ribose and the sulfamate moiety, and the main chain carbonyl of Gly316. In the second conformation, observed with the cross-linked luciferase enzyme, the side chain amine of Lys443 adopts a nearly identical position as Lys529 (Figure 2). Additionally, from the C-terminal domain, Gln448 also rotates into the binding pock-et where it interacts with a sulfamate oxygen. These interac-tions may help stabilize the new C-terminal conformation.

His245 is part of the A4 motif that exhibits distinct side chain torsional rotations in the two conformational states1. In the adenylate-forming conformation, this residue points toward the active site, preventing access of water and perhaps stabilizing the approach of the negatively

charged luciferyl carboxylate and ATP. In the thioester-forming members of the enzyme family, the rotation of the C-terminal domain withdraws the A4 aromatic side chain out of the active site, clearing a tunnel for CoA or pantetheine to approach the adenylate intermediate. This side chain motion appears to be induced by a residue from the C-terminal domain10. In the Ppy 9-C108/C447 luciferase structure, electron density for the side chain of His245 is weak, suggesting the His245 side chain is adopting both side chain torsional conformations. On the basis of comparisons to other ANL enzymes, the residue from the C-terminal domain that would normally stabilize the rotated His245 position is Tyr447, one of the residues that was mutated to cysteine for the cross-link formation. The missing side chain of Tyr447 most likely prevents the full rotation of His245 and contributes to the torsional disorder seen in the crystal structure.

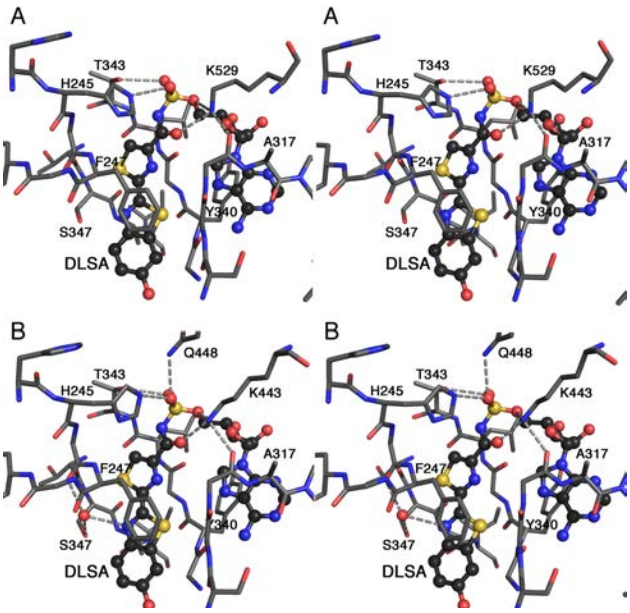
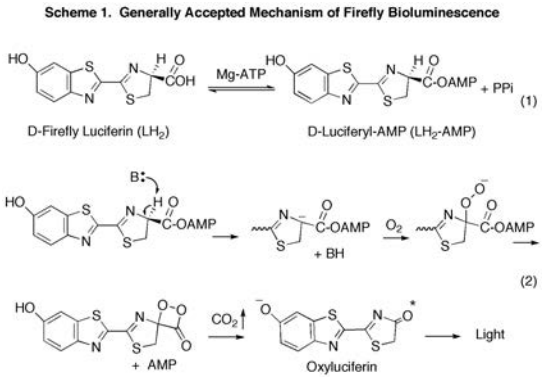


Figure 2. Stereoview of the active site of luciferase complexes with DLSA in the A. adenylate-forming and B. second catalytic conformation observed with the cross-linked enzyme.

Insights into the biochemical mechanism of LH2-AMP oxidation. The structure provides insights into several interesting features of the luciferase-catalyzed reaction. Like the homologous acyl-CoA synthetases, luciferase can catalyze thioester formation between CoA and luciferin, using LH2-AMP as an intermediate[18]. In the cross-linked luciferase, the pantetheine tunnel seen in other ANL enzymes is intact. At the end of this tunnel is the C4 carbon of DLSA. Luciferase contains in the A8 motif a conserved glycine residue, Gly446, that lines this tunnel. This glycine forms a distorted β-sheet interaction with the amide nitrogens of the pantethene moiety and likely facilitates CoA binding in this alternate reaction. A G446I mutation in luciferase specifically impairs the oxidative reaction[13] sug-gesting that O2 approaches the intermediate through this same tunnel. Side chain rotation of His245, as discussed above, further expands the tunnel for access to C4 (see Figure S1 of Supporting Information).

The structure also provides an explanation for the stereo-chemical requirements of luciferase. It is reported that while the enzyme is able to catalyze thioester-formation between CoA and both



D- and L-luciferin, only D-luciferin is able to support the generation of light[19]. This observation is explained by the accessibility of the carbonyl of the adenylate of both isoforms to the pantetheine tunnel to be attacked by the CoA thiol. In contrast, the C4 proton must be available for removal, possibly as a hydrogen atom (see below), in order to initiate the oxidative reaction required for the generation of light. The structure demonstrates that the adenylate with L-luciferin would project the C4 hydrogen toward the core of the protein and away from the pantetheine tunnel, preventing proton or hydrogen atom abstraction.

The generally accepted mechanism of firefly luciferase-catalyzed light production, as advanced primarily by the work of E. H. White and co-workers[20], is initiated by abstraction of the C4 proton of LH2-AMP by a presumed active site nucleophile to produce a carbanion (Scheme 1). The involvement of a peroxy anion/hydroperoxide intermediate leading to a highly reactive dioxetanone intermediate, while not directly observed, is very reasonable and consistent with the general current view of excited state formation in bioluminescence[21]. It is, however, problematic that peroxide formation from unactivated molecular O2 is a spin forbidden process[22]. Perhaps an even more serious shortcoming of an enzyme mechanism requiring C4 carbanion formation is the absence of a required nucleophile proximal to C4 (Figure 2B). While it might be tempting to suggest that the side chain imidazole of H245 is the critical residue, our prior finding[23] that the H245A and H245F luciferase mutants retained ~20% of the activity of the wild-type enzyme make it quite unlikely. The absence of an active site base near C4 certainly does not disprove this mechanism; however, it may be worthwhile to consider alternatives that do not require carbanion formation. We therefore propose for consideration another mechanism for formation of the key LH2-AMP hydroperoxide intermediate that involves O2 abstraction of a H atom producing C4 and hydroperoxide radicals as the initiating and rate determining step (Scheme 2A). This proposal is discussed further in the Supporting Material. Additional mutagenesis and chemical model studies are in progress to evaluate the radical-based mechanism.

REFERENCES

(1) Gulick, A. M. ACS Chem Biol 2009, 4, 811-827.
(2) Marahiel, M. A.; Stachelhaus, T.; Mootz, H. D. Chem Rev 1997, 97, 2651-2674.
(3) Gulick, A. M.; Starai, V. J.; Horswill, A. R.; Homick, K. M.; Escalante-Semerena, J. C. Biochemistry 2003, 42, 2866-2873.
(4) Hughes, A. J.; Keatinge-Clay, A. Chem Biol 2011, 18, 165-176.
(5) Kochan, G.; Pilka, E. S.; von Delft, F.; Oppermann, U.; Yue, W. W. J Mol Biol 2009, 388, 997-1008.
(6) Reger, A. S.; Wu, R.; Dunaway-Mariano, D.; Gulick, A. M. Biochemistry 2008, 47, 8016-8025.
(7) Mitchell, C. A.; Shi, C.; Aldrich, C. C.; Gulick, A. M. Biochemistry 2012, 51, 3252-3263.

(8) Sundlov, J. A.; Shi, C.; Wilson, D. J.; Aldrich, C. C.; Gulick, A. M. Chem Biol 2012, 19, 188-198.
(9) Reger, A. S.; Carney, J. M.; Gulick, A. M. Biochemistry 2007, 46, 6536-6546.
(10) Wu, R.; Cao, J.; Lu, X.; Reger, A. S.; Gulick, A. M.; Dunaway-Mariano, D. Biochemistry 2008, 47, 8026-8039.
(11) Wu, R.; Reger, A. S.; Lu, X.; Gulick, A. M.; Dunaway-Mariano, D. Biochemistry 2009, 48, 4115-4125.
(12) Branchini, B. R.; Murtiashaw, M. H.; Magyar, R. A.; Anderson, S. M. Biochemistry 2000, 39, 5433-5440.
(13) Branchini, B. R.; Southworth, T. L.; Murtiashaw, M. H.; Wilkinson, S. R.; Khattak, N. F.; Rosenberg, J. C.; Zimmer, M. Biochemistry 2005, 44, 1385-1393.
(14) Branchini, B. R.; Rosenberg, J. C.; Fontaine, D. M.; Southworth, T. L.; Behney, C. E.; Uzasci, L. J Am Chem Soc 2011, 133, 11088-11091.
(15) Nakatsu, T.; Ichiyama, S.; Hiratake, J.; Saldanha, A.; Kobashi, N.; Sakata, K.; Kato, H. Nature 2006, 440, 372-376.
(16) Branchini, B. R.; Murtiashaw, M. H.; Carmody, J. N.; Mygatt, E. E.; Southworth, T. L. Bioorg Med Chem Lett 2005, 15, 3860-3864.
(17) Conti, E.; Franks, N. P.; Brick, P. Structure 1996, 4, 287-298.
(18) Fraga, H.; Fernandes, D.; Fontes, R.; Esteves da Silva, J. C. FEBS J 2005, 272, 5206-5216.
(19) Nakamura, M.; Maki, S.; Amano, Y.; Ohkita, Y.; Niwa, K.; Hirano, T.; Ohmiya, Y.; Niwa, H. Biochem Biophys Res Commun 2005, 331, 471-475.
(20) White, E. H.; Rapaport, E.; Seliger, H. H.; Hopkins, T. A. Bioorg Chem 1971, 1, 92-122.
(21) Navizet, I.; Liu, Y. J.; Ferre, N.; Roca-Sanjuan, D.; Lindh, R. Chemphyschem : a European journal of chemical physics and physical chemistry 2011, 12, 3064-3076.
(22) Min, C. G.; Ren, A. M.; Li, X. N.; Guo, J. F.; Zou, L. Y.; Sun, Y.; Goddard, J. D.; Sun, C. C. Chem Phys Lett 2011, 506, 269-275.
(23) Branchini, B. R.; Magyar, R. A.; Murtiashaw, M. H.; Anderson, S. M.; Zimmer, M. Biochemistry 1998, 37, 15311-15319.

Protein production, crystallization, and structure determination of luciferase in adenylate-forming conformation. Several structures of firefly luciferase have been determined. To best compare the structure of the chemically cross-linked luciferase, we first determined the structure of the wild-type enzyme in the adenylate-forming conformation bound to DLSA (Figure 1A). Wild-type luciferase was produced and purified as described previously1,2. The enzyme (13 mg/mL in 25 mM Tris-Cl pH 7.8 (at 4 °C), 1 mM DTT, 1 mM EDTA and 200 mM ammonium sulfate) was incubated with a 8.5-fold molar excess of the adenylate analog 5'-O-[N-dehydroluciferyl)- sulfamoyl]adenosine (DLSA)3 and frozen in small aliquots for crystallization experiments. Prior to crystallization experiments, the protein was incubated with 0.8 mM CoA. Crystals of the wild- type enzyme were grown by hanging drop vapor diffusion using 1:1 mixtures of protein and a crystallization cocktail composed of 300mM Na/K Tartrate, 20% PEG 6000, and 100mM Tris pH7.5. Crystals were cryo-protected in the crystallization cocktail with a stepwise gradient of 8- 23% ethylene glycol.

Crystals of the wild-type luciferase enzyme were used for data collection at SSRL beamline 9-1. The diffraction data were processed with iMOSFLM4 and POINTLESS and SCALA from the CCP4 suite of programs5. Analysis indicated partial twinning in the trigonal space group P3121 and refinement was performed with twinning modules of both REFMAC6 and PHENIX7. The structure was solved by molecular replacement using the structure of the *L. cruciata* luciferase enzyme8 as a search model. This model was chosen because it contained the C-terminal domain in the expected adenylate-forming conformation rather than the open conformation of the original luciferase structure9. The final model contains residues 4-541 and 1-550 for chains A and B, respectively. Electron density was clear for the DLSA ligand; there was no evidence of CoA binding in the crystal lattice. Both chains have disordered loops at residues 199-203, the conserved A3 motif involved in binding the phosphates of ATP, while chain A contains two additional disordered loops at residues 460-465 and 489-493 of the C-terminal domain. The two chains superimpose with a root mean square (rms) displacement of Ca atoms of 0.6 Å over 522 residues. Because of the improved completeness of chain B, all figures were prepared with chain B.

Protein production, crystallization, and structure determination of cross-linked luciferase. The Ppy 9- C108/C447 protein used to obtain the structure of the cross-linked enzyme contains several amino acid substitutions10. The substitutions T214A, A215L, I232A, F295L, and E354K create a more stable protein. The mutations C82S, C216A, C258S, and C391S changed the endogenous cysteine residues to prevent non-specific reaction with the chemical cross-linker. Finally, Tyr447 and Ile108 were both mutated to cysteine residues to enable the cross-link with 1,2-bis(maleimido)ethane (BMOE) to form. The luciferase protein was expressed, purified, and cross-linked as described previously10. The chemically cross-linked Ppy 9- C108/C447 protein (11 mg/mL in 25 mM Tris-Cl pH 7.6 (at 4 °C), 1 mM EDTA and 200 mM ammonium sulfate) wasincubated with a 6-fold molar excess of DLSA and stored as frozen aliquots at -80°C. The protein was thawed for crystallization daily. Crystals of the cross-linked enzyme were grown by hanging drop vapor diffusion using 1:1 mixtures of protein and a crystallization cocktail composed of 50 mM NaCl, 30% PEG 4000, and 50 mM HEPPS pH 8.5. Crystals were cryo-protected by quickly soaking in the reservoir solution containing 75 mM NaCl and 10 % 2,3-Butanediol.

The cross-linked luciferase enzyme incubated with the DLSA inhibitor crystallized as large diamonds in a space group shown to be different than trigonal crystals of the wild-type enzyme. An initial anisotropic 2.7 Å resolution data set was collected at SSRL beamline 9-2 and used to solve an initial structure by molecular replacement using the N-terminal domain of 1LCI, the wild-type luciferase structure9. Electron density for the C-terminus was visible and was manually modeled. Weak density for the DSLA and the cross-linker were visible but were not present at full occupancy.

An improved dataset was collected from a second crystal that was grown under the same conditions but was cryo-protected in the crystallization drop through the addition of 1µL of 50% PEG 4000. This crystal was isomorphous to the previous crystal and diffracted to 2.4Å. The protein atom coordinates from the earlier dataset refinement were used as a search model for molecular replacement with PHASER that resulted in excellent electron density and a crystallographic R-factor of 26% (Rfree of 34%). The model was then completed through iterative manual model building and refinement with both REFMAC5 and PHENIX. During refinement, the DLSA ligand was included. The final model contains residues 4-543 of chain A and 1-543 of chain B. Each chain contains the disordered loop of the conserved nucleotide binding motif at Gly200 through Leu204. Additionally, the A chain contains weak density for residues Ile231 through Ala236, which was included in the final model, and at the A10 motif, Thr527 and Gly528, which was not of sufficient quality to justify including these residues.

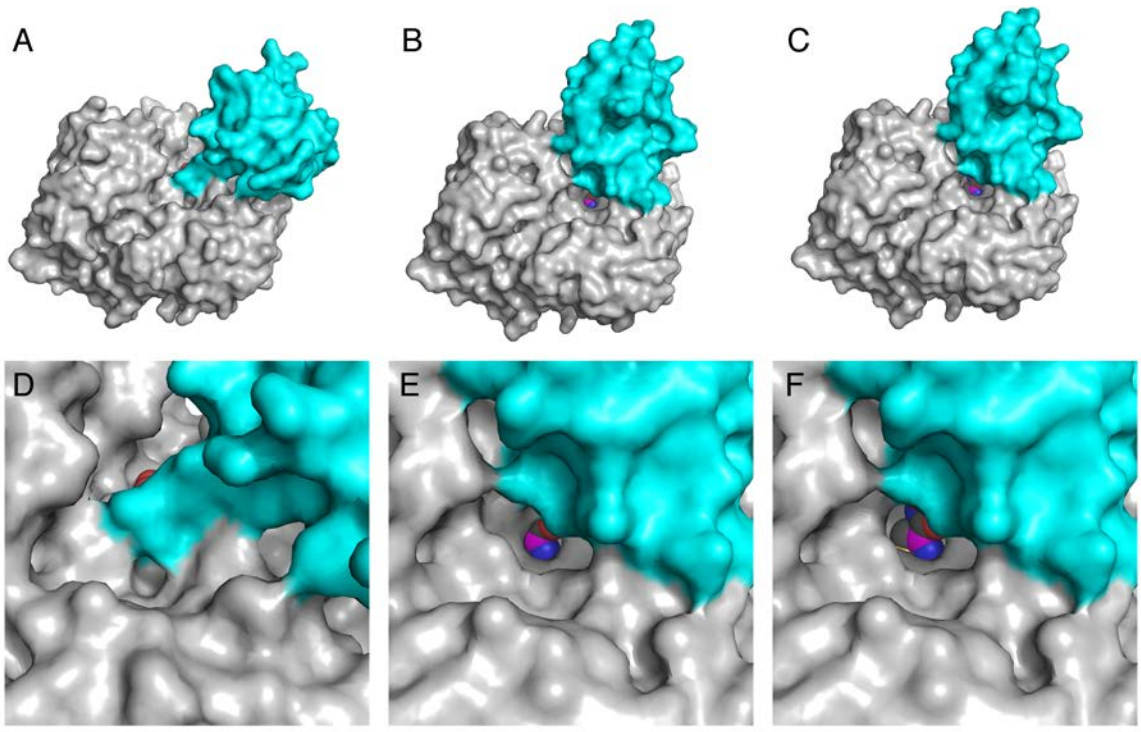
Electron density of the form *Fo-Fc* displayed at 2.5σ was continuous between the side chains of Cys108 and Cy447 in Chain B and nearly continuous in chain A. The succinimide rings are hydrolyzed and the resulting carboxymethyl groups are disordered and are not apparent in the density.
The final data collection and refinement statistics are present in Table 1. The structure factors and atomic coordinates have been deposited in the Protein Data Bank (4G36, wild-type luciferase with DLSA, and **4G37,** the Ppy 9- C108/C447 mutant protein with DLSA and the chemical cross-linker).

Table S1. Crystallographic Data Collection and Refinement Statistics.

Data Collection
Beamline Wavelength (Å) Space GroupUnit cell a, b, c (Å) Resolution range No. Obs.
No. unique reflections Completeness % (outer) I/σ (outer)Rmerge (%) (outer)
Structure Refinement
Resolution range (Å)
R-factor (%) (outer shell)
Rfree (%) (outer shell)
No. protein/solvent atoms
No. Ligand/Crosslinker/ion atoms
RMSD bond distances (Å)
RMSD bond angles (°)
Ramachandran Favored and allowed (%)

Molprobity clashscore (percentile)
Wild-type Luciferase
SSRL 9-1 0.9795P312193.0, 93.0, 297.6 40.0 – 2.62 Å 296866
45585 99.6 (99.6) 15.6 (5.9) 7.9 (27.7)
40 – 2.62 18.8 (23.9) 23.8 (27.5) 8069 / 72 80 / 0 / 0 0.008
1.17 93.4, 5.8
18.3 (80%)
Cross-linked Luciferase
SSRL 9-20.9795C222175.6, 184.1, 170.5 30.0 – 2.4 Å 161696
44431 94.4 (82.3) 15.3 (3.8) 6.7 (22.8)
30–2.4 20.0 (25.5) 26.0 (33.6) 8083 / 190 80/20/20 0.0081.1696.5, 3.0
10.0 (94 %)

Analysis of the Pantetheine Tunnel in Cross-linked Luciferase.
Biochemical data show that luciferase is able to use CoA as a substrate to react with the LH2-AMP adenylate to form LH2-CoA. The structure of the cross-linked luciferase demonstrates a pantetheine tunnel, similar to what has been observed with the ANL enzymes (Figure S1). The tunnel leads directly to the C4 carbon, which is the site of nucleophillic attack by the CoA thiol, or of proton or hydrogen abstraction for the light producing reaction.
Figure S1. View down the pantetheine tunnel. The luciferase protein is shown in full view (A-C) and zoomed closer on the pantetheine tunnel (D-F) in the same orientation. Wild-type luciferase is shown in the adenylate-forming conformation (A and D). The Ppy 9- C108/C447 luciferase is shown in the second, oxidative conformation (B and E). Panels C and F show the same Ppy 9-C108/C447 structure with His245 rotated through the side chain torsion angle from $\chi 1 = -167°$ to –67° to further clear the pantetheine tunnel, as observed in the ANL adenylating enzymes. The DLSA molecule is shown with grey carbons, blue nitrogens, red oxygens, and yellow sulfurs. The C4 carbon atom is highlighted in magenta.



An alternate mechanism for formation of the hydroperoxide intermediate.
We propose for consideration an alternative mechanism for the formation of the key LH2-AMP hydroperoxide intermediate that involves O2 abstraction of a H atom producing C4 and hydroperoxide radicals as the initiating and rate determining step (Scheme S1A). Recombination of these radicals would then provide the LH2-AMP hydroperoxide required for dioxetanone formation. Lys443 is likely significant in stabilizing/adopting the second conformation (Figure 2B). It may also be important for maintaining a conformation that allows O2 access to the C4 position of LH2-AMP. While a radical mechanism is speculative, one potentially important feature is that the radical, but not the anion, can be resonance stabilized through the entire aromatic ring system (Scheme S1B), perhaps facilitating the proposed role for O2. Another attractive feature of a radical-based mechanism is that it readily accounts for the formation of dehydroluciferyl-AMP and H2O2, which are well-documented11 side products that form during the luciferase-catalyzed oxidation of LH2-AMP (Scheme S1C).

Scheme S1. Radical-Based Concepts of Firefly Reactions